

А.Д. Иванников, В.П. Кулагин,  
А.Н. Тихонов, В.Я. Цветков

# ГЕОИНФОРМАТИКА

А.Д. Иванников, В.П. Кулагин,  
А.Н. Тихонов, В.Я. Цветков

ГЕОИНФОРМАТИКА

Москва  
МАКС Пресс  
2001

**А.Д. Иванников, В.П. Кулагин,  
А.Н. Тихонов, В.Я. Цветков**

# **ГЕОИНФОРМАТИКА**

**Москва  
МАКС Пресс  
2001**

УДК 681.518(075.32)

ББК 32.973.202

Г35

**Геоинформатика/** А.Д. Иванников, В.П. Кулагин,  
Г35 А.Н. Тихонов, В.Я. Цветков. – М.: МАКС Пресс, 2001. –  
349 с.  
ISBN 5-317-00310-5

В книге излагаются основы одной из самых молодых наук – геоинформатики. Методы геоинформатики находят широкое применение во всех сферах человеческой деятельности, поэтому книга представляет интерес для специалистов разных специальностей и направлений.

Она предназначена для преподавателей вузов, студентов и аспирантов, а также для специалистов, применяющих методы геоинформатики в бизнесе, управлении и проектировании.

УДК 681.518(075.32)

ББК 32.973.202

**ISBN 5-317-00310-5**

© А.Д. Иванников, В.П., Кулагин,  
А.Н. Тихонов, В.Я. Цветков, 2001

## ВВЕДЕНИЕ

Геоинформатика является наукой, изучающей все аспекты сбора, обработки и представления информации о свойствах объектов, процессов и явлений, происходящих на Земле.

Как и всякая наука геоинформатика имеет основную область исследований и основной метод исследований. Определение области и метода содержится в названии геоинформатики. Термин "геоинформатика" состоит из двух частей "гео" и "информатика".

Первая часть "гео" определяет область исследования науки - объекты и явления, происходящие на земной поверхности. В геоинформатике используют пространственно-временные данные, в которых содержится информация о пространственном положении объектов, их свойствах и времени, для которого эти свойства имели место.

"Информатика" в составе термина "геоинформатика" определяет основной метод исследования: объектов исследования изучаются на основе компьютерных технологий, при этом привлекаются данные математики, картографии, геодезии и других наук, занимается изучением и развитием систем сбора, передачи, обработки и хранения информации с помощью автоматизированных методов обработки и автоматизированных систем.

Начинала свое развитие информатика (геоинформатика также) как прикладная инженерная дисциплина. Возможно по этой причине у многих людей старшего поколения с информатикой долгое время ассоциировалось изучение либо языков программирования и основ математической логики, либо решение прикладных задач при помощи компьютера.

В настоящее время исследования в таких областях, как социология и психология уже немыслимы без информационных систем и технологий.

Современный мир формирует новое мировоззрение, в котором осознается сущность информационных процессов в их неразрывной связи с энергией и материей. Особенностью информатизации на данном этапе является создание информационных пространств в рамках отдельных организаций, отраслях или страны [6, 25, 26, 28, 61, 82-85].

Информатика становится метанаукой для самых различных областей, играя важную роль в развитии современного общества. Геоинформатика также является метанаукой, ибо она многоаспектна в смысле приложений и включает множество разделов.

Обычно выделяются два основных аспекта геоинформатики: научный и технологический аспект. Научный аспект связан с разработкой концепций, теоретических основ, методов моделирования, организацией моделей и структур данных.

Технологический аспект геоинформатики связан с обработкой и передачей информации с помощью компьютерных технологий.

Из этого следует, что геоинформатика является своеобразным синтезом наук, занимающимися изучением процессов и явлений, происходящих на Земле.

Таким образом, можно определить геоинформатику как науку, занимающуюся изучением пространственно-временных процессов, явлений и систем в масштабах, сопоставимых с объектами на земной поверхности. В область геоинформатики не попадают явления микромира и макромира. Методы геоинформатики применимы не только на земной поверхности, но и для изучения других планет.

На рис. 1. геоинформатика показана в виде модели дерева. Корни дерева составляют основообразующие науки. Ствол - объединяющие науки. Ветви и листья образуют приложения геоинформатики.

К числу основообразующих наук относятся геодезия, картография, география, фотограмметрия. Геодезия - это основа точных расчетов и измерений. Картографическая информация является основной формой представления объектов геоинформатики. Дистанционное зондирование и его технологии составляют основу получения данных в современной геоинформатике.

Другую группу наук составляют науки, позволяющие формализовать и строить информационные модели. Это теория множеств, теория графов, топология и др.

Еще один блок составляют науки, связанные с информационными системами: теория баз данных, программирование, компьютерная графика и др. Следующий блок составляют науки, обеспечивающие анализ в геоинформатике: системный анализ, структурный анализ, моделирование, теория принятия решений.

Объединяющими науками являются новые информационные технологии, теория автоматизированных информационных систем. Все это осуществляется в рамках необходимости защиты информации и информационной безопасности.

Приложения геоинформатики разнообразны. Простейшее приложение - это автоматизированное картографирование; более сложное - исследование природных ресурсов.



Рис.1. Дерево геоинформатики

Приложениями геоинформатики являются различные виды анализа. К ним относятся: статистический, экономический, пространственный, транспортный, анализ сетей.

Земельные информационные и земельные кадастровые системы являются важным аспектом приложения геоинформатики. Геомаркетинг, логистика и риелтерская деятельность являются бизнес-приложениями геоинформатики.

Блок приложений управления включает муниципальное, региональное, отраслевое и государственное управление, управление транспортными сетями и т.д.

Блок мониторинга состоит из экологического мониторинга, военной разведки, анализа чрезвычайных ситуаций и др.

Это далеко не полный перечень приложений геоинформатики призван дать представление о возможностях и особенностях этой науки.

Именно многоаспектность геоинформатики позволяет утверждать, что это фундаментальная наука, объединяющая гуманитарные и технические направления.

Геоинформатика как и информатика состоит из частей [58, 73, 104]. Но если в информатике выделяют две составляющие "научную" и прикладную, то геоинформатику следует рассматривать как совокупность трех частей.

Основными частями геоинформатики являются "общая геоинформатика", "специальная геоинформатика", "прикладная геоинформатика".

**1. Общая геоинформатика** это раздел геоинформатики, занимающийся исследованием и разработкой научных основ, концепций, обобщенным анализом геоинформатики и геоинформационных систем безотносительно к их прикладному характеру. Эта часть геоинформатики базируется на: теории множеств, общей теории систем, математической логике, структурном и лингвистическом анализе, теории моделирования и построения абстрактных моделей данных, географии, геодезии, картографии, фотограмметрии, классификационном анализе и т.д.

Раздел общей геоинформатики может быть определен как фундаментальный. В этой части геоинформатики представлены научный и технологический аспекты, но большее внимание уделено играет научному аспекту.

**2. Прикладная геоинформатика** изучает практические методы работ с геоинформационными системами и геоинформационны-

ми технологиями. В этой части информатики представлены научный и технологический аспекты, но более детально рассматривается технологический аспект: именно в этой части геоинформатики изучают специализированные системы обработки пространственно-временных данных, называемые геоинформационными системами (ГИС).

Кроме того, в данном разделе геоинформатики изучают технологии сбора, хранения, обновления, обработки и представления данных. Эти технологии называют геоинформационными технологиями.

**3. Специальная геоинформатика** служит основой для дополнительного анализа систем и методов обработки пространственно-временных данных, и оценки качества. Эта часть геоинформатики связана с теорией информации, анализом качества информации, вопросами стандартизации разработки геоинформационных систем и продуктов, вопросами тестирования информационных средств и продуктов, вопросами защиты информации, вопросами телекоммуникационного обмена, вопросами преобразования видов и форматов данных в геоинформатике и т.д.

Большой раздел геоинформатике посвящен изучению автоматизированных информационных и телекоммуникационных систем. Одними из основных информационных систем в геоинформатике являются геоинформационные системы (ГИС). Это обобщенное понятие, поскольку функциональное назначение ГИС может быть различным.

Их практической реализацией могут быть разные системы: геодезические информационные системы (ГИС), геологические информационные системы (ГИС), географические информационные системы (ГИС), земельные информационные системы (ЗИС), автоматизированные системы картографирования (АСК) и т.д.

На основе ГИС создают системы экологического мониторинга, маркетинговые информационные системы, экономические информационные системы, статистические информационные системы, системы пространственных сетей (транспорт, коммуникации) и т. д. [8, 11, 31, 44 -48, 51, 52, 56, 68, 71, 72, 76, 81, 89, 97, 105, 107, 110, 112, 120, 132].

В геоинформатике промежуточной формой хранения и обработки данных являются цифровые модели.

Термин "цифровые" заимствован из названия "цифровые вычислительные машины" (ЦВМ) [8, 10, 20, 91, 97, 108]. На ранних



этапах развития вычислительной техники существовало два типа вычислительных машин аналоговые и цифровые.

Аналоговые обрабатывали непрерывные аналоговые сигналы. Они использовали электронные схемы, осуществляющие функциональное преобразование информации в соответствии с заранее заданным алгоритмом. По этой причине они были узко специализированными.

Цифровые вычислительные машины использовали дискретную информацию. Для того, чтобы обработать информацию с их помощью она должна быть дигитализована (digital - цифра). Эти вычислительные машины были более универсальны, так как позволяли обрабатывать данные с помощью набора программ.

В настоящее время аналоговые вычислительные машины практически не используются за исключением специальных устройств анализа данных. Все персональные компьютеры являются цифровыми и поэтому данный термин не употребляют по отношению к компьютерам.

Термин "цифровой" сохранился как характеристика для некоторых данных и систем (цифровые карты, цифровые снимки, цифровые фотокамеры, цифровая фотограмметрия). В настоящее время он означает, что информация в этих данных и системах содержится в дискретной форме и предназначена для обработки с помощью современных компьютерных технологий.

Термин "Цифровая модель местности" используется для наименования информации о местности, которая хранится в дискретных наборах данных.

Технологии обработки данных в геоинформатике называют геоинформационными технологиями. Они включают технологии работы как с геоинформационными системами, так и с другими системами (базы данных, GPS), которые применяют в геоинформатике для сбора, хранения, обработки и анализа информации.

Итак, обобщая все вышесказанное, геоинформатика это развивающаяся наука, объединяющая многие дисциплины, в число которых входят геодезия, картография, вычислительная техника, география, фотограмметрия, дистанционное зондирование, статистика, и другие отрасли знаний, имеющие дело с обработкой и анализом пространственно локализованных данных.

Следует отметить еще одно сходство геоинформатики и информатики. Оно заключается в том, что цели обучения данным дисциплинам весьма схожи.

На первых этапах (70 -80-е годы) подготовка специалистов в области информатики была направлена на обучение программированию.

В (90-годы) основной целью обучения информатике было подготовка специалистов по обработке информации.

В настоящее время основной целью изучения информатики является подготовка специалистов, свободно владеющими компьютерными и телекоммуникационными технологиями в любой предметной области.

Подобная основная цель обучения стоит и перед геоинформатикой - подготовить специалистов, свободно владеющих и использующих геоинформационные технологии для решения любых практических задач (от составления карт до бизнеса, экономики и управления) в любой технической или гуманитарной области.

## *Общая геоинформатика*

### **1. Информационные ресурсы и модели**

#### **1.1. От информации к информационным ресурсам**

Информатизация и раст все более заметную роль в науке, образовании, развитии производства, управлении, бизнесе. Можно утверждать, что кризис терпят в первую очередь организации, в которых информационной деятельности и информационным технологиям не уделяется должного внимания и средств.

В условиях информатизации общества термин “информация” породил целый ряд производных понятий: информационная система, информационная основа, информационная культура, информационный бизнес, информационная этика, информационное общество и др.

Как следствие информатизации многие научные дисциплины выделили соответствующие направления с присоединением слова “информационный”: информационная экономика, информационная география, информационная медицина, информационное право и др.

Появилось научное направление, определившее новую науку - геоинформатику.

Переход отдельных стран и всего человечества в информационное общество означает, что большая часть населения будет работать в сфере производства информации и информационных услуг и значительная часть общества будет являться потребителем информационных продуктов и услуг.

В этих условиях возрастает значение применения методов управления и производства, основанных на информационных технологиях и информационных ресурсах

Если термин “информация” достаточно определен, то термин “информационные ресурсы” объясняют по разному.

Для анализа эволюции информации от совокупности данных к информационным ресурсам необходимо проследить основные этапы информатизации общества.

Информатизация общества происходила не только в последние десятилетия, когда ее достижения стали особенно заметны. Если социальные и экономические потрясения замечают и ощущают большинство членов общества, то информационные изменения заметны, как правило, специалистам, причем только тем кто связан с

анализом и обработкой информационных процессов. Наиболее заметными этапами информатизации являются информационные революции.

Информационные революции происходят в масштабах человечества. В процессе развития человечества прошло несколько информационных революций.

**Первая** информационная революция связана с изобретением письменности, что позволило формализовать и осуществлять передачу знаний и накопленного опыта человечества на тот период развития.

Основное направление информационных технологий сводилось к передаче и накоплению (архивированию) знаний. Производительность и формализация знаний были несовершенны. Письменностью владели единицы (посвященные).

**Вторая** (XVI в.) определена изобретением книгопечатания. В сущности это была первая информационная технология, которая радикально повлияла на развитие цивилизации и передачу информации. Производительность накопления и передачи знаний резко увеличилась.

Этот период характерен тем, что образование и письменность стали проникать в широкие слои общества и не быть уделом избранных. Основное направление информационных технологий было направлено на упорядочение функционирования общественной жизни, разработку идей преобразования. Возникла цензура как способ контроля информационных потоков общества со стороны власти.

**Третья** (конец XIX в) стала возможной благодаря открытию электричества и созданию таких технических средств как телефон, радио, телеграф и др. При этом возникли новые технологии: передачи данных, кодировки, декодирования, преобразования одних видов информации в другие. В этот период появилась возможность преобразования разных видов информации к единой форме и развитию на этой основе формальных методов анализа данных.

**Четвертая** (70 гг. XX в.) связана с изобретением микропроцессорной техники и персональных компьютеров. Она взаимосвязана с появлением информационного общества как очередного этапа развития наиболее передовых стран. Этот период характерен развитием систем искусственного интеллекта (ИИ), как новой основы развития общества.

Следует подчеркнуть, что изобретение компьютера не привело к информационной революции, а только предшествовало ей. Дли-

тельное время доступ и работа с компьютером были уделом избранных специалистов, как и во времена первой информационной революции, когда уделом избранных было владение письменностью.

Данное положение привело к закономерному противоречию: уровень использования информационных средств на производстве, в науке не соответствовал реальной потребности общества. Только после появления устройств массового пользования типа персональных компьютеров общество получило широкий доступ к информационным технологиям, что обусловило качественно новый подход к управлению бизнесом и производством.

Для последней информационной революции характерно качественное изменение информации. Из средства передачи знаний или данных информация становится ресурсом производства, средством его изменения, товаром. Следствием этого процесса становится появление новых информационных технологий (НИТ) и информационного бизнеса.

Информационный бизнес это бизнес, связанный с информацией не только как с товаром, но и, в первую очередь, как с ресурсом производства или средством улучшения коммерческой или предпринимательской деятельности.

Ресурсы являются источником любого производства. **Ресурсами производства** называют совокупность природных, материальных, социальных факторов, которые могут быть использованы в процессе создания товаров, услуг и прочих ценностей.

В экономических теориях ресурсы принято делить на четыре категории:

- **природные** - потенциально пригодные для производства естественные силы природы и вещества. Служат основой производства;
- **материальные** - все созданные человеком средства производства, которые сами являются результатом производства. Служат основой производства;
- **трудовые** - население в трудоспособном возрасте как средство создания материальных благ. Этот вид ресурсов обычно оценивают по трем параметрам: социально-демографическому, профессионально-квалификационному, культурно-образовательному. Служат основой и способствуют созданию производства;
- **финансовые** - денежные средства выделяемые на организацию производства Служат основой производства.

С одной стороны информационные ресурсы как *совокупность данных* относятся к материальным ресурсам. С другой стороны информационные ресурсы как *средство искусственного интеллекта* включаются в трудовые ресурсы. Это определяет двойственное значение информационных ресурсов и предопределяет выделение их в отдельную группу.

Под информационными ресурсами (с позиций НИТ) понимают совокупность знаний, данных, технологий, систем искусственного интеллекта, систем анализа и обработки информации, *служащих основой* или *способствующих* созданию материальных благ или различных видов продукции.

Результатом использования информационных ресурсов в бизнесе и производстве являются информационные продукты и услуги, информационные технологии и программные продукты.

Информационным продуктом будем называть совокупность данных, сформированную производителем в вещественной или невещественной форме. Это могут быть технологии, программы, наборы данных в различных формах представления, базы данных, экспертные системы и т.д.

Информационной услугой назовем получение и предоставление в распоряжение пользователя информационных продуктов.

Одним из основных информационных продуктов является программная продукция (ПП). В настоящее время рынок программного обеспечения включает следующие основные группы ПП:

- обработка текста (редакторы, процессоры, переводчики);
- электронные таблицы;
- системы управления базами данных (СУБД);
- программы компьютерной графики (включая деловую);
- геоинформационные системы (ГИС);
- редакционно-издательские системы;
- компьютерные игры;
- языковые средства;
- интерфейсы и пр.

Как всякая другая продукция информационная должна отвечать определенным стандартам, требованиям и критериям. Для нее существует понятие качества и нормы его контроля.

Основными видами информационных программных средств, используемых в информационном бизнесе, являются: информационная основа, информационные программно-технологические средства, информационные (включая программные) продукты и услуги.

В настоящее время в связи с расширением сфер информационного бизнеса более широко применяют понятие **информационный продукт** которое является обобщающим для **информационного продукта** и **программного продукта**.

Информация может быть рассмотрена как совокупность данных. Информационные ресурсы содержат некий интеллектуальный потенциал.

Интеллектуальный потенциал основан на использовании неких моделей, способных к принятию решений. Такие модели могут содержать более простые модели данных.

Понятие "информационной модели" как некоторого целенаправленного формализованного отображения существующей системы информации лежит в основе большинства используемых методов обработки в информационных системах.

Информационная модель обеспечивает формализованное представление (алгебраическое, графическое и др.) используемых данных и их взаимосвязей и содержит несколько уровней описания: предметный, связанный с областью обработки информации; системный, связанный с методов организации и способами обработки; базовый, определяемый выбором базовых моделей данных, независящих от области применения информационной модели.

Переход от информации к информационным ресурсам означает переход от совокупности разрозненных данных к совокупности взаимосвязанных моделей, обладающих свойством ресурсности. Это отличает модели данных от моделей информационных ресурсов.

**Ресурсность** модели заключается в свойстве накопления информации (опыта) и возможности использования накопленной информации для изменения своих возможностей. Ресурсность модели означает возможность повышения качества информации и расширяет возможность применения модели.

Основными характеристиками качества информации наряду с общепринятыми для других видов продукции являются: репрезентативность, содержательность, прагматизм, достаточность, точность, актуальность, устойчивость. Те же самые характеристики применимы для информационных ресурсов.

Ресурсность информационных моделей позволяет по мере накопления данных (информации) улучшать качество информационных ресурсов.

Таким образом, качественное повышение информации возможно за счет организации ее как ресурса, т.е. организации как совокуп-

ности моделей обладающих свойством накопления и использования данных для улучшения решений, получаемых с помощью этих моделей.

## 1.2. Модели информационных ресурсов

В настоящее время в связи с широкой информатизацией общества появляются новые понятия, такие как информационные модели, информационные ресурсы, информационная инфраструктура и др. Следует отметить, что само понятие компьютерных и информационных технологий изменилось и требует уточнения.

В конце 60-х-начале 70-х годов основное направление компьютерных технологий было направлено на совершенствование методов обработки информации, в первую очередь алгоритмической. Наибольшее внимание уделялось системам обработки данных (СОД). Основным типом моделей в системах обработки информации были информационные модели, представляющие наборы данных.

К началу 80-х годов развитие компьютерных технологий и применение компьютерной техники меняется. Главным действующим лицом становится не программист, а пользователь – специалист в области использования и обработки информации. При этом методика обработки информации от алгоритмов прямого счета смещается в сторону эвристической обработки в человеко- машинных системах (ЧМС). Это обусловило интенсивную разработку методов и средств, обеспечивающих эффективное взаимодействие пользователя и компьютера.

Направление работ компьютерных технологий было направлено на совершенствование методов накопления информации, на создание информационных систем, а не только систем обработки. Возрастает важность систем обработки данных и управления (СОДУ), содержащих не только систему обработки, но и систему хранения информации.

Появляются интегрированные информационные системы, меняющие концепцию однокомпонентной обработки информации в сторону многоаспектной многокомпонентной обработки данных. С появлением интегрированных систем появилось новая модель данных – *интегрированная информационная основа*.

Именно на этом этапе актуализировалось понятие информационных ресурсов – ресурсов доступных не только специалистам в области программирования или обработки информации, но и широкому кругу массового пользователя. В этот период развития ин-



форматики появляется новый тип моделей - информационно-ресурсные модели.

Благодаря интеграции и интегрированным системам к началу 90-х годов появилась качественно новая реальная возможность разработки методов и средств искусственного интеллекта (ИИ). Новая методика позволила обеспечить эффективное взаимодействие пользователя и компьютерной системы. Это привело к расширению и развитию нового класса моделей – интеллектуальных моделей.

В настоящее время под интеллектуализацией компьютерных технологий понимают два направления их развития. Первое направление, называемое (smart) – «сообразительным», можно условно отнести к интеллектуальным. Оно направлено на развитие интерфейса и упрощения процесса общения человека и компьютера, постоянного расширения доли машины в совместной с человеком деятельности по решению задачи, выдачи подсказок и рекомендаций в критических или тупиковых ситуациях. Однако решение в этих технологиях остается за человеком.

Второе направление основано на том, что главное решение принимает компьютер, а человек лишь контролирует процесс обработки. Наибольшее применение такой подход получил в системах проектирования сложных объектов, в современных системах поддержки принятия решений (OLAP).

Таким образом, развитие компьютерных технологий привело к расширению понятия информационная модель. Появилось три класса информационных моделей: информационно-описательные, информационно-ресурсные, интеллектуальные.

1. К информационно-описательному (дескриптивному) классу относят модели, построенные как описание некоего процесса, явления, объекта, сущности и т.д. Эти модели выполняют функции информационного сообщения. Они модели могут быть простыми, составными и пр. Примером таких моделей могут служить файл, текстовый документ, речевое сообщение.

2. Информационно-ресурсные модели способны аккумулировать данные для собственного улучшения и оптимизации. Примером информационно-ресурсных моделей могут быть модели базы данных, человеческая память.

3. Интеллектуальные модели - это класс моделей, обладающих способностью к накоплению информации, самосовершенствованию и осуществлению действий независимо от субъекта или объекта, создавшего эти модели. Примером этих моделей могут быть базы

знаний, некоторые типы компьютерных вирусов, модели реакции человека на воздействие внешней среды.

Все три класса моделей относят к информационным ресурсам. Они имеют общие признаки, так как являются развитием класса информационных моделей, и различия.

Первый класс информационных моделей получил наиболее широкое распространение в неавтоматизированных технологиях. В информационных технологиях наибольшее значение получил информационно-ресурсный класс моделей.

Второй класс моделей (информационно-ресурсные) в отличие от первого обладает способностью к самосовершенствованию при поступлении новой информации. Он должен содержать некий механизм, связывающий разрозненные данные, что упорядочивает их хранение, анализ и обновление. Использование этих моделей позволяет получать новые информационные совокупности.

Сходство первого и второго класса моделей заключается в их "пассивности". Это означает, что всякое их изменение (обновление, актуализация) осуществляется при внешнем воздействии на модель.

Третий класс моделей в отличие от первых двух является активным. Модели данного класса наиболее совершенны из числа информационно-ресурсных моделей. Они дополняются механизмом накопления знаний, позволяющим осуществлять активные действия без соответствующих запросов к этим моделям.

Процесс преобразования информационно-описательной модели к информационно-ресурсной показан на рисунке 1.1.

Информационно-описательная модель представляет собой описание наборов данных. Накопление информации в этой модели ведет к суммированию данных либо к замене одних данных другими. В этой модели данные хранятся в виде независимых наборов.

Однако эту независимость можно снять.

Над данными, входящими в описательную информационную модель возможно выполнение ряда действий:

анализ и выделение связей внутри и между наборами данных.

классификация всех данных.

определение отношений между данными.

Совокупность этих действий позволяет преобразовать описание наборов данных в систему связанных данных. Эта система создает информационно-ресурсную модель.

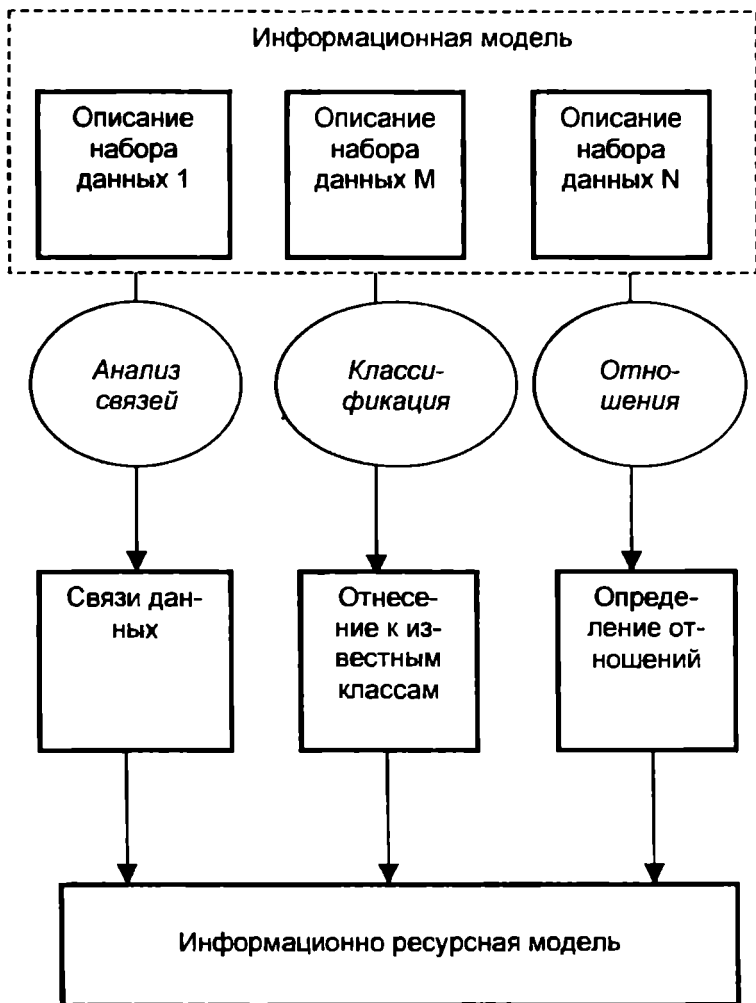


Рис.1.1. Связь между информационно-описательной и информационно-ресурсной моделью.

Система связанных данных имеет некий механизм, образованный отношениями и классификацией данных. Этот механизм позволяет строить новые конструкции данных на основе имеющихся.

Механизм этой системы позволяет ее оптимизировать, по мере поступления новой информации в систему данных в тех случаях, когда в ней заранее определены параметры оптимизации.

В этом и заключается *ресурсность* модели, т.е. возможность получения новых данных и возможность совершенствования модели на основе внутреннего механизма без специальной обработки.

### 1.3. Информационное моделирование как метод познания

Информатизация становится все более заметным явлением в развитии общества. В условиях информатизации общества от каждого члена общества требуется умение оперировать различными видами и формами информации. Наряду с понятиями "информационная система", "информационная основа" сосуществуют понятия "информационная культура", "информационный бизнес", "информационная этика", "информационное моделирование" и др.

Переход отдельных стран и всего человечества в информационное общество означает, что большая часть населения будет работать в сфере производства информации и информационных услуг, а значительная часть общества будет являться потребителем информационных продуктов и услуг.

Рассматривая глубинный процесс информатизации общества – интеграцию, можно говорить о том, что информатика является основой интеграции информационного общества. В этих условиях информационное моделирование становится ключом к освоению и познанию окружающего мира.

Для полного отображения объектов реального мира и всех их свойств понадобилась бы бесконечно большая база данных. Использование процедур абстракции сводит множество данных к конечному объему. Это достигается применением моделей, сохраняющих основные свойства объектов исследования и не содержащих второстепенные свойства.

В основе большинства используемых методов обработки в информационных системах и технологиях лежит понятие информационной модели – некоторого целенаправленного формализованного отображения существующих объектов реальности. Информационная модель служит описанием и может быть построена для объек-

та, совокупности объектов, сложной системы, информационной системы.

Информационная модель системы (объекта) обеспечивает формализованное представление исследуемых элементов системы и их взаимосвязи и содержит разные уровни описания: предметный, системный, базовый.

Для построения оптимальной информационной модели, которая отвечает целям повышения эффективности и качества информационной системы, необходимо определить основные понятия и конструкции, характерные для предметных областей.

Информационное моделирование заключается в умении построения информационных моделей и их анализа для изучения объектов реального мира. Информационное моделирование можно рассматривать как современную информационную технологию. Оно включает умение создавать, интерпретировать и моделировать различные информационные модели.

Информационное моделирование начинается с предварительной структуризации предметной области: объекты реального мира подвергаются классификации, фиксируется совокупность подлежащих отображению в базе данных типов объектов.

Напомним, что предметная область представляет собой структурированную совокупность информационных компонентов, отображающую рассматриваемую сферу деятельности информационного объекта. Предметная область считается определенной в тех случаях, когда известны существующие в ней объекты, их свойства и отношения.

Современное информационное моделирование требует умения работать с информацией не только как с данными, но и осуществлять качественный переход от информационных описательных моделей к ресурсным, от ресурсных к интеллектуальным. По этой причине можно утверждать, что информационное моделирование есть основа будущей интеллектуализации общества, связующее звено между информационным и интеллектуальным обществом.

Информационное моделирование использует следующие понятия:

- базовые категории;
- классификации;
- концепция отношений;
- системный подход;
- структурный анализ;

- построение информационных единиц

**Базовыми категориями** являются *объект моделирования* и *метод моделирования*.

**Объектом моделирования** может быть объект окружающего пространства, модель объекта, набор данных, система, процесс, решение задачи, прогнозная оценка и т.п.

Для каждого типа объектов выявляется совокупность свойств. Посредством этих свойств описываются конкретные объекты этого типа в базе данных и виды отношений (взаимосвязей) между этими объектами. Затем решаются вопросы о том, какая информация об объектах должна быть представлена в базе данных и как ее представить с помощью данных.

**Метод моделирования** определяется набором допустимых условий и правил осуществления преобразований над объектами моделирования.

**Классификация** означает определение классов (подклассов, групп) моделей и преобразований над ними, определение свойств классов и ограничений.

**Концепция отношений** означает определение и выбор классов отношений между объектами. Отношения могут быть различными: пространственные, временные, функциональные, логические, вероятностные, организационные, количественные, качественные. Они отражают многоаспектность объектов и связей между ними. Отдельные отношения становятся доминирующими в конкретных предметных областях. Например, в геоинформатике доминирующими становятся пространственно-временные отношения.

**Системный подход** основан на абстрактном рассмотрении объектов моделирования и связей между их элементами на основе обобщенного использования понятия "система". Этим понятием может быть обозначен, объект, процесс, набор данных, модель. При системном анализе выявляются наиболее общие части структуры, связи и отношения между элементами "системы". Все это можно определить понятием элементы системы. Выявленные элементы служат основой дальнейшей детализации "системы".

Системный подход дает обобщенное представление объекта моделирования в виде системы. Дальнейшую детализацию осуществляет структурный анализ.

**Структурный анализ** основан на рассмотрении элементов объектов моделирования на основе функционального описания. Таким образом, структурный анализ означает переход с уровня абст-

ракий на уровень функций. Структурный анализ позволяет представить структуру объекта моделирования как совокупность функциональных блоков преобразующих информационные модели.

В отличие от системного анализа, оперирующего с абстрагированными данными, структурный оперирует с реальными данными. Он требует сведения разнообразных реальных данных и информационных потоков в единую систему. Такая единая система, называемая *информационной основой* [5], требует выбора информационных единиц.

Информационные единицы, выбранные в процессе информационного моделирования, должны подчиняться определенным правилам - синтаксису.

Таким образом, информационное моделирование в самом общем виде, независимо от области применения должно отвечать определенным концепциям, должно быть направлено на отображение и изучение окружающей действительности.

Содержание информационного моделирования можно показать с помощью парадигматических отношений.

Процесс исследования объектов окружающего мира можно упрощенно представить в виде следующих отношений

**объект → содержание → отображение**

При использовании информационного моделирования этот процесс интерпретируется таким образом

**сущность → описание свойств → отнесение свойств к известным классам → описание отношений между свойствами → абстрактное описание системы свойств и отношений → функциональное описание свойств и отношений → выбор наборов информационных единиц → изменение одних (входных) наборов информационных единиц → исследование изменения других (выходных) наборов информационных единиц**

В геоинформатике термин "свойство" часто заменяют термином "атрибут".

Следует подчеркнуть, что информационное моделирование значительно эффективнее, если оно использует в качестве основной информационно-ресурсную модель.

Информационно-ресурсные модели имеют более длительный

жизненный цикл по сравнению с информационно-описательными за счет механизма самосовершенствования. Это обеспечивает преемственность и долгий жизненный цикл информационного моделирования в условиях смены технических и программных средств.

#### 1.4. Семиотический подход к организации моделей данных

Одним из назначений моделей данных является реализация информационного обмена. Примерами стандартов информационного обмена могут служить естественный язык и современные языки (включая правила их употребления), которые сформировались как результат требований некоего стандартного обмена информацией.

В технических науках примером моделей информационного обмена служат различные формализованные языки.

Таким образом, информационные средства и технологии используют де-факто стандарты, которые официально такими не называются, но, тем не менее, такими являются. Это говорит о важности и необходимости исследования моделей как систем информационного обмена.

Исследованием обмена информацией на основе системы знаков, образующих формальные языки, занимается *семиотика*.

Семиотика - это теория языковых знаков, связей их с друг другом, связей знаков с человеческим мышлением, связей с объективной реальностью и человеком. Данная теория является общей для всех языков, независимо от их словарного состава, грамматик и способов их возникновения.

Понятием знака семиотика охватывает не только такие языки как русский, английский, немецкий и т.д., но другие образования, такие как математические символы, знаки уличного движения, условные знаки на топографических картах, знаки различия военнослужащих и т.п.

Переходя на формальное обозначение, следует выделить четыре фактора, которые учитываются семиотикой:

Z - знак;

A - мысленные образы, отражения (модели);

O - объекты отражения;

M - людей или их мышление.

Различные взаимосвязи перечисленных факторов определяют различные разделы семиотики. Кратко охарактеризуем

**Синтактика** - раздел семиотики, который абстрагируется от



всех факторов за исключением знака. Она исследует связи между знаками некоторого языка. Синтактика устанавливает правила построения составных языков, создает критерии определения принадлежности совокупности знаков к определенному языку.

**Семантика** (от греч. семантикос - обозначающий, сема- знак) - раздел семиотики, изучающий интерпретацию знаковой системы. Он исследует отношения между знаками  $Z$  и отражениями  $A$ , связи между словами языка и соответствующим им понятиям. Знаки  $Z$  являются формой существования мысленных образов  $A$ , последние являются значениями знаков  $Z$ . Так семантика рассматривает отношения между знаками языка и их значениями.

**Сигматика** изучает отношения между  $Z$  и  $O$ . Языковые значения  $Z$  - это имена обозначения объектов  $O$ .

**Прагматика** - раздел семиотики, изучающий отношение использующего знаковую систему к самой знаковой системе.

Опишем как выглядит применительно к информатике и геоинформатике семиотический анализ информационной единицы (модели, системы, процесса, объекта, технологии). Модель должна включать три составные части:

**синтаксис** - правила построения и критерии принадлежности к виду;

**семантику** - методы отображения информации объекта в информацию модели и саму эту информацию;

**прагматику** - полезность, методы оценки полезности модели.

Таким образом, семиотика выделяет три основные характеристики модели как знаковой системы: синтаксис, семантика и прагматика.

Этот подход определяет содержание элементарной (атомарной) модели как элементарной единицы данных, включающей правила построения, имя объекта, свойства объекта, значения свойств.

## 2. Основные виды и характеристики моделей

### 2.1. Сильно и слабо типизированные модели

Одно из противоречий описания моделей заключается в том, что с одной стороны необходимо отразить общие свойства класса объектов, с другой - индивидуальные признаки конкретного объекта. Этот аспект рассмотрения приводит к разделению моделей данных на два класса: сильно типизированные и слабо типизированные.

**Сильно типизированные** - это модели, в которых большинство данных могут быть отнесены к узкому подклассу (типу). Примером сильно типизированных данных в геоинформатике служат координатные (метрические) данные. Другим примером может служить столбец данных в таблице.

Если исходные данные нельзя отнести к одному типу, то их можно с помощью искусственных приемов (введением дополнительных условий или ограничений) разделить и отнести к набору типов.

**Слабо типизированными** моделями называют те, в которых данные разнородны по формату, структуре. Они слабо связаны условиями относительно известных типов. Примером слабо типизированных моделей могут быть текстовые описания.

Сильно типизированные модели применяют для обработки. Слабо типизированные модели используют для обобщенных описаний.

## 2.2. Статические и динамические модели

С позиций изменчивости можно выделить два класса моделей: статические и динамические.

**Статическими** моделями называют модели, неизменяющиеся с течением времени. Они служат для описания процессов и явлений, не зависящих от времени.

**Динамические** модели - это модели, которые допускают изменение параметров и структур во времени и служат для описания изменения процессов и моделей именно во времени.

Построение динамических моделей (например, для задач управления), как правило, более сложно, чем построение статических.

Поэтому в некоторых случаях применяют квазидинамические модели как упрощение динамических.

**Квазидинамические** модели - это модели, в которых временной интервал действия модели разбивается на периоды, для каждого из которых строится статическая модель. Таким образом, квазидинамические модели можно рассматривать как совокупность меняющихся и взаимосвязанных статических моделей.

Примерами динамических и статических моделей в геоинформатике могут служить два вида электронных карт.

Электронные карты в режиме разделения времени (электронные атласы) представляют реализацию статических моделей, в то

время как электронные карты в реальном масштабе времени могут служить примером динамической модели.

Следует подчеркнуть, что понятие изменчивости моделей данных в информационных системах - относительно, так как вся информация носит временной характер и через какой-то период времени требует обновления (актуализации).

Из этого следует, что применение понятий статистические и динамические модели данных требует указания периода времени, который используется в процессе исследований или указания альтернативной модели при сравнении с исходной.

### **2.3. Аналоговые и дискретные модели**

По способу отображения объекта модели могут быть аналоговыми либо дискретными. Примерами таких моделей могут служить обычный фотоснимок и сканированное изображение снимка.

Аналоговые модели в свою очередь разбиваются на две группы: прямой и косвенной аналогии.

К первой группе относятся модели, создаваемые на основе физического моделирования: аналоговые карты, модели судов, самолетов, гидротехнические сооружения и т. п.

Во вторую группу (модели косвенной аналогии) входят модели создаваемые на основе математического моделирования (аналитического описания), например, цифровая модель рельефа, построенная на основе аналитического описания поверхности.

Дискретные модели основаны на замене непрерывных функций набором дискретных значений аргументов и функций. Дискретность определяется *шагом квантования*. При необходимости сохранения информативности дискретной модели по отношению к объекту шаг квантования выбирается с учетом теоремы Шеннона - Котельникова.

### **2.4. Масштаб действия модели**

Для моделей, описывающих пространственные объекты или явления, существует понятие масштаба действия модели. Масштаб действия модели определяет область действия и применимость модели. Например, модель экологического мониторинга городского района неприменима для описания экологической ситуации города или области. Модель экологической ситуации региона может иметь меньшую детализацию, чем модель города.

Масштаб действия модели определяется функциями и террито-

риальными размерами объекта исследований и числом характеристик модели, называемым также пространством параметров.

## 2.5. Жизненный цикл модели

Жизненный цикл (ЖЦ) модели является важной характеристикой, которая определяет период возможного применения модели. ЖЦ – временная характеристика любой продукции. Жизненный цикл модели определяется периодом, в течение которого модель можно эффективно использовать [36]. Поэтому при разработке любой модели необходимо оценить ее жизненный цикл как семиотическую составляющую, определяющую полезность модели.

Жизненный цикл включает разные периоды разработки и применения модели. Можно выделить шесть фаз жизненного цикла информационной модели:

- фаза разработки концепции;
- фаза разработки проекта модели;
- фаза реализации модели;
- фаза эксплуатации и актуализации модели
- фаза модернизации;
- фаза завершения эксплуатации модели

На первой фазе создается *концептуальная* схема разработки и использования модели.

**Концептуальная схема** определяется как непротиворечивая совокупность высказываний истинных для данной предметной области, включая возможные состояния, классификации, законы, правила [11].

На этапе разработки проекта осуществляется структурное проектирование, декомпозиция задач и поиск решений задач с помощью данной модели.

На фазе реализации осуществляется построение модели на основе программно-аппаратных средств.

На фазе эксплуатации осуществляют практические работы и актуализацию данных для приведения характеристик модели в соответствии с изменяющимися внешними условиями.

Модернизация модели направлена на ее совершенствование и продление периода эксплуатации.

Фаза завершения эксплуатации модели очень важна, так как помимо консервации модели, которая уже не достаточно отражает существующую реальность, необходимо обеспечить преемственность в передаче данных от устаревшей модели к новой.

## 2.6. Формы представления моделей данных

Аспект представления моделей разграничивает модели по четырем основным формам представления: аналитические, табличные, графические и графовые. Эти формы представления моделей показаны на рис. 1.2.

**Аналитическая форма** (рис. 1.2. а) представляет модель в виде формулы, аналитического выражения, совокупности аналитических выражений (уравнений). Она применяется при известных законах поведения модели или объекта.

**Графическая форма** (рис. 1.2. б) использует отображение совокупности моделей или данных в виде кривых, графиков, диаграмм.

Она наиболее часто применяется при наличии статистических данных и при известном аналитическом описании модели, то есть является формой взаимосвязанной с аналитической.

**Табличная форма** (рис. 1.2. в) дает представление модели или ее характеристик в виде одной или совокупности взаимосвязанных таблиц. Она применяется при описании атрибутов и при сборе статистической информации.

Следует подчеркнуть, что при этом данные в ячейках таблицы не могут заноситься произвольно, они подчиняются определенным правилам, в частности, по столбцам располагают типизированные данные.

**Графовая форма** (рис. 1.2. г) основана на представлении модели в виде графической схемы называемой графом. Она применяется при описании структур моделей данных, процессов обработки или управления и описании сложных систем.

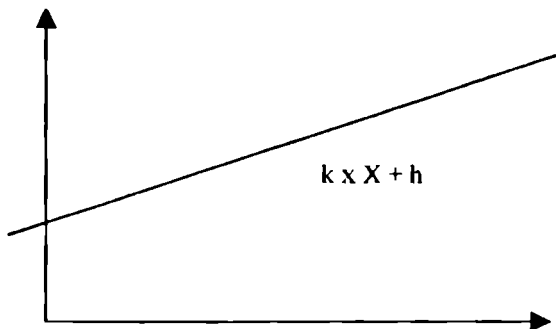
Схема включает элементы графа, называемые вершинами (узлами) и ребрами (дугами). В отличие от произвольно нарисованной схемы графовая модель, как и табличная модель, строится по определенным правилам.

В частности, каждое ребро может быть ориентировано, если определен путь от одной вершины к другой, и не ориентировано, что соответствует возможному пути от одной вершины к другой в обоих направлениях. Простейшим примером неориентированного графа может быть схема метрополитена.

Кроме вершин и ребер существуют другие элементы, значение которых не столь существенно на данном этапе рассмотрения моделей.

$$Y = k \times X + h$$

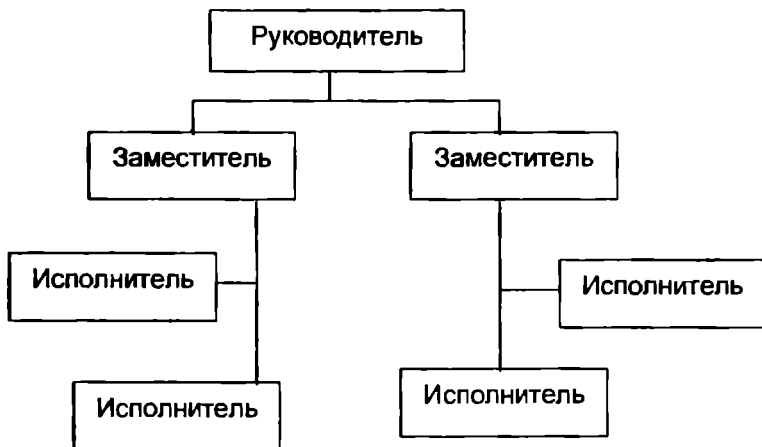
А) Аналитическая форма



Б) Графическая форма

N	X	Y
1	0	h
2	1/k	h+1

В) Табличная форма



Г) Графовая форма

Рис. 1.2. Формы представления моделей

В целом формы представления моделей реализуются средствами компьютерной графики и деловой графики.

Перечисленные формы могут быть связаны между собой. Как видно на рис. 1.2. аналитическая форма может быть переведена в графическую форму. Набор значений аргументов позволяет по аналитической форме получать набор значений функций. Эти наборы можно поместить в таблицу и образовать табличную форму представления этой же функции. Таким образом, имеется прямая связь между аналитической, графической и табличной формами моделей.

Более сложная связь существует между графовой и табличной формами. В теории моделирования и теории баз данных известна процедура нормализации, которая позволяет связывать иерархическую и табличные модели [56].

### 3. Базовые модели данных, используемые в геоинформатике

#### 3.1. Иерархическая модель

Одной из форм представления моделей является графовая, которая позволяет изображать их в виде схем. Эта форма применяется для отображения структуры: систем, объектов, процессов или явлений. Модели, которые позволяют отображать некую структуру, называют структурными.

К наиболее простым *структурно определенным* относится иерархическая модель (рис.3.1). Она моделирует структуру обычного дерева [35, 30, 56].

Граф иерархической модели (ее схемное представление) включает два типа элементов: дуги и узлы (записи, вершины).

Дуги соединяют только те узлы, между которыми существует функциональная связь. Количество дуг, выходящих из узла, определяет тип связи, (характеристику, называемую также степенью ветвления).

В иерархической модели данных связи между ее частями являются жесткими, а ее структурная диаграмма должна быть упорядоченным деревом. Это означает, что иерархическая модель напоминает структуру дерева.

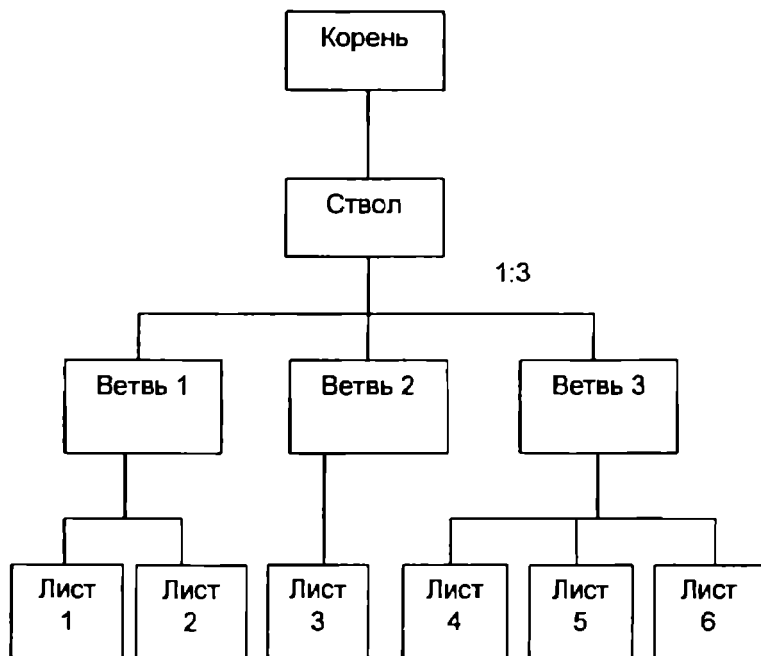


Рис.3.1. Пример обобщенной иерархической модели

На рисунке 3.1. приведена обобщенная схема иерархической модели, содержащая четыре уровня. Цифрами показаны степени ветвления для узлов. Для описания разных уровней применяют следующие понятия : *корень*, *ствол*, *ветви*, *листья*, что подчеркивает сходство структуры иерархической модели со структурой дерева.

Дуги должны быть всегда направлены от корня в листья дерева, они всегда ориентированы. Между двумя узлами в иерархической модели может быть не более одной дуги. Говорят, что дуга исходит из типа *родительского* узла и заходит в тип *дочернего* узла.

Примером организационной иерархической модели служит модель, приведенная на рис.1.2. г).

Одно из важных понятий иерархической модели - уровень. Уровень определяет совокупность равных по функциональному зна-



чению узлов. Так на рисунке 1.2. г) верхний уровень соответствует руководителю. Более низкий уровень определяет положение заместителей. Еще более низкий уровень определяет исполнителей и т.д.

Таким образом, главная задача иерархической модели состоит в выделении основных частей и демонстрации подчиненности одних частей модели другим.

На рисунке 3.2. приведен другой пример иерархической модели. Он представляет собой известный всем пользователям компьютерных технологий фрагмент дерева каталогов. Это схема размещения информации на жестком диске. Она появляется при использовании опции "Tree" - "дерево" в программах FAR, Norton Commander или им подобных.

Особенностью иерархической модели, приведенной на рисунке 3.2. является то, что дуги идут не сверху вниз или снизу вверх, а слева направо. Такая модель носит название **Е-дерево**.

**Иерархический путь** - это последовательность узлов, начинающаяся от корневого. Он определяется как длина пути от корневой записи, выраженная в числе дуг. Так, тип корневой записи "диск" находится на нулевом уровне, "корневой каталог" - на первом, "подкаталог" - на втором, имя файла - на третьем и т.д.

В иерархических моделях связь между узлами должна иметь вид аналогичный структуре дерева. Это означает, что узлы должны разветвляться и не должны соединяться. В теории моделирования такую связь обозначают  $1 : n$  ("один ко многим"). Данное условие является ограничением.

Недостатком иерархических моделей является увеличение времени поиска при большом числе уровней, поэтому на практике не используют модели при большом числе уровней (более 10). Это явилось причиной того, что появившиеся первоначально иерархические базы данных уступили место более быстрым реляционным.

В то же время иерархические модели эффективно применяют для составления классификаторов.

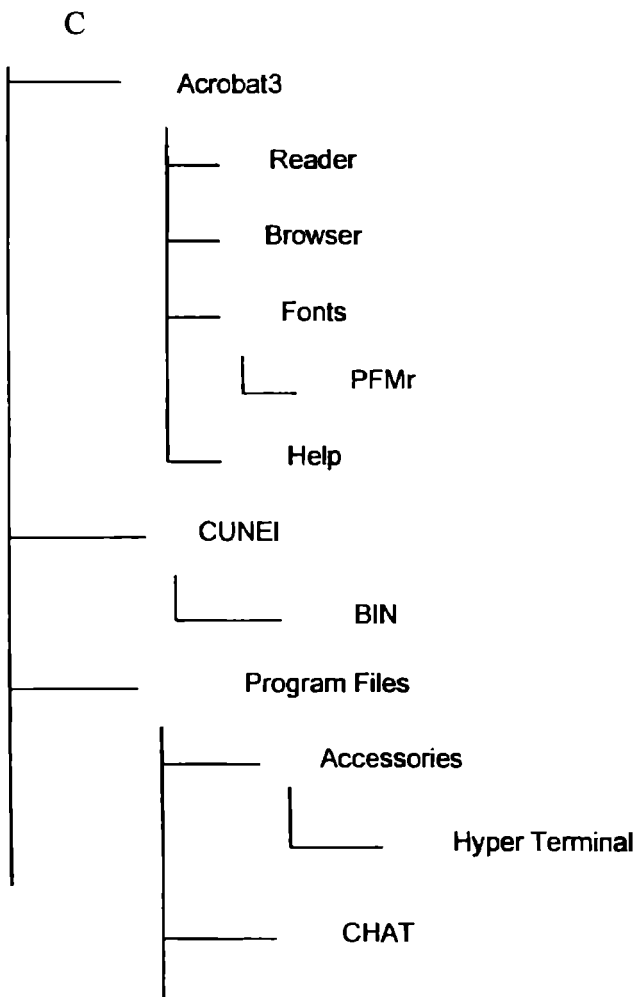


Рис. 3.2. Иерархическая модель дерева каталогов

### 3.2. Квадратомическое дерево

Модель квадратомическое дерево является частным случаем иерархической модели [30].

**Квадратомическое дерево** - это иерархическая модель с постоянной связью между ее элементами 1 : 4.

Квадратомическое дерево является характерной моделью области геоинформатики. Ее впервые применили в теории обработки изображений, откуда была заимствована в геоинформатику.

Целью применения данной модели явилась необходимость кодирования растрового изображения и обеспечения устойчивой связи между его элементами. При использовании квадратомического дерева двумерная геометрическая область рекурсивно подразделяется на квадранты, что определило название данной модели.

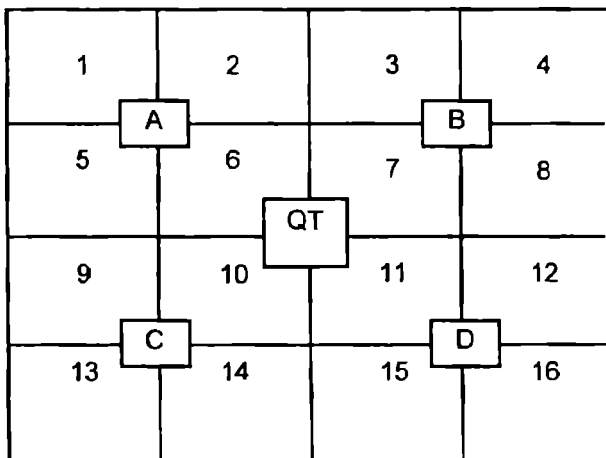


Рис. 3.3 Исходное изображение

На рис. 3.3. показан фрагмент изображения - двумерная область QT, состоящей из 16 пикселей. Каждый пиксель обозначен цифрой. Вся область разбивается на четыре квадранта: А, В, С, D. Каждый из четырех квадрантов является узлом квадратомического дерева.

Большой квадрант  $QT$  становится корневым узлом квадратомиического дерева, а меньшие квадранты появляются на более низких уровнях.

Технология построения квадратомиического дерева основана на рекурсивном разделении квадрата на квадранты и подквадранты до тех пор, пока все подквадранты не станут однородными по отношению к значению изображения (цвета) или пока не будет достигнут предопределенный заранее наименьший уровень разрешения.

Если регион состоит из  $2n \times 2n$  пикселей, то он полностью представлен на уровне  $n$ , а единичные пиксели находятся на нулевом уровне. Квадрант уровня  $l$  ( $0 < l < n$ ) содержит  $2^l \times 2^l$  пикселей, всего  $4^l$ .

Исходное изображение можно рассматривать как объект, который структурируется на части для последующего построения модели с организацией связей между частями.

На рис. 3.4 показано квадратомиическое дерево, построенное по исходному изображению, приведенному на рис. 3.3.

Структура модели, приведенная на рис.3.4 являет собой классический пример E-дерева. Преимущество такой структуры состоит в том, что регулярное деление обеспечивает накопление, восстановление и обработку данных простым и эффективным способом. Простота проистекает из геометрической регулярности разбиения, а эффективность достигается за счет хранения только узлов с данными, которые представляют интерес.

Модели, основанные на квадратомиических деревьях, обеспечивают расчеты площадей, распознавание образов, выявление связанных компонентов, определение соседства, преобразование расстояний, деление изображений, сглаживание данных и усиление краевых эффектов

Следует отметить, что особенность такого построения обеспечивает возможность автоматизированного кодирования, автоматизированного хранения и автоматизированной обработки графической информации.

Эта модель допускает возможность параллельной обработки, что существенно повышает производительность и эффективность при больших объемах информации, что является характерной особенностью изображений.

Уровень 2      Уровень 1    Уровень 0

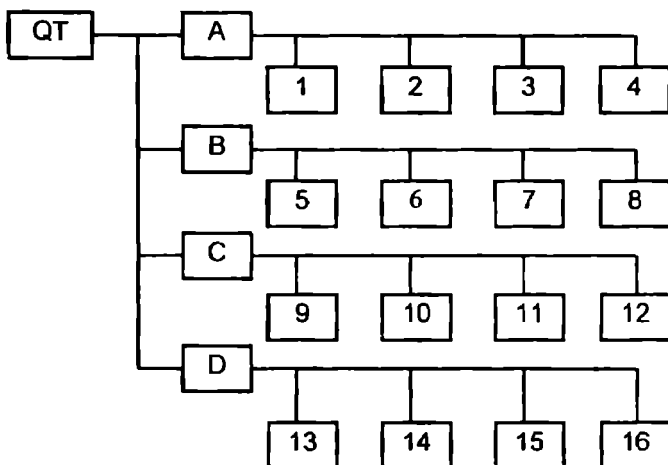


Рис.3.4. Организация связи между элементами изображения в виде квадратомиического дерева

В приведенном на рис. 3.3 примере результатом построения модели на рис.3.4. явилась трехуровневая модель. На этой модели наглядно видно, что число узлов на уровнях может быть различно, но для квадратомиического дерева это число не произвольно, а зависит от номера уровня.

Нумерация уровней может вестись от корня к листьям или наоборот, что на практике встречается реже.

### 3.3. Реляционная модель

В современных информационных системах и базах данных наиболее широко представлены реляционные модели (РМ). Их называют еще табличными, потому, что информация в таких моделях хранится в виде таблиц.

Название *реляционная* модель происходит от английского слова (relation), что означает отношение.

Реляционная модель данных, разработанная Коддом еще в 1969-70 гг. на основе математической теории отношений, опирается на систему понятий, важнейшими из которых являются: таблица, отношение, строка, столбец, первичный ключ, внешний ключ, домен.

Доменом называется совокупность значений, не повторяющихся в одном столбце.

Таблица имеет имя уникальное внутри базы данных. Таблица имеет столбцы и строки. Каждый столбец - это атрибут. Каждая строка - это запись или кортеж. Совокупность атрибутов в таблице задает отношение между ними.

Первая строка таблицы содержит набор атрибутов. Остальные строки содержат значения атрибутов

Каждый столбец имеет имя, которое записывается в верхней части таблицы. Оно должно быть уникальным в таблице, однако различные таблицы могут иметь столбцы с одинаковыми именами

Столбцы расположены в таблице в соответствии с порядком следования их имен при ее создании. В отличие от столбцов строки не имеют имен, порядок их следования в таблице не определен, а количество логически не ограничено.

Так как строки в таблице не упорядочены, то в таблице вводят один или несколько столбцов, значения в которых однозначно идентифицируют каждую ее строку. Такой столбец (или комбинация столбцов) называется **первичным ключом** (primary key).

Взаимосвязь таблиц - важнейший элемент реляционной модели данных. Она поддерживается **внешними ключами** (foreign key).

Для обработки данных, размещенных в таблицах, нужны дополнительные сведения о данных, например описатели таблиц, столбцов и т.д. Их называют обычно **метаданными**. Метаданные также представлены в табличной форме и хранятся в словаре данных (data dictionary).

Данные информационной системы должны быть однозначными и непротиворечивыми. В таком случае говорят, что реляционная модель удовлетворяет условию целостности (*integrity*). При этом на реляционную модель накладываются некоторые ограничения, которые называют ограничениями целостности (*data integrity constraints*).

Существует несколько типов ограничений целостности. Например, требуется, чтобы значения в столбце таблицы выбирались только из соответствующего домена. На практике учитывают и более сложные ограничения целостности, в частности, целостность по ссылкам (*reference integrity*). Ее суть заключается в том, что внешний ключ не может быть указателем на несуществующую строку в таблице.

Ключевой столбец	Название атрибута 1	Название атрибута 2	Название атрибута 3	Название атрибута 4
1	Ко	р	те	ж
2	1	1	1	1
3	2	1	2	2
4	3	2	3	2
5	4	3	4	1
6	2	2	5	1
7	3	3	6	2

Рис.3.5. Реляционная модель

На рисунке 3.5. приведен пример реляционной модели. По существу это обычная табличная форма со строками и столбцами. Ключевой столбец является первичным ключом. Верхнюю строку таблицы образуют названия атрибутов, а последующие строки значения (знаки) атрибутов. Кортеж - совокупность значений по строке, связанных отношениями.

Домен для первого атрибута рис.3.5. составляют значения 1, 2, 3, 4. Домен для второго атрибута составляют значения 1, 2, 3. Домен для третьего атрибута составляют значения 1, 2, 3, 4, 5, 6. Домен для четвертого атрибута составляют значения 1, 2.

### 3.4. Инфологическая модель данных "Сущность-связь"

Модель данных "сущность-связь" или ER-модель (Entity Relationship Model) дает представление о предметной области в виде объектов, называемых сущностями, между которыми фиксируются связи.

Для каждой связи определено число связываемых ею объектов. На схеме сущности изображаются прямоугольниками, связи - ромбами. Число связываемых объектов указывается цифрой на линии соединения объекта и связи.

Появление моделей данных типа "сущность-связь" было обусловлено практическими потребностями проектирования баз данных для коммерческих СУБД. Такие модели имеют много общего с иерархическими и сетевыми моделями данных.

Теоретической основой этого подхода является известная модель, введенная М. Ченом в 1976 г. и получившая широкое распространение в качестве средств концептуального проектирования баз данных.

В основе модели Чена лежит представление о том, что предметная область состоит из отдельных объектов, находящихся друг с другом в определенных связях. Объекты описываются различными параметрами или атрибутами; однотипные объекты описываются одним и тем же набором параметров и объединяются во множества или классы (сущности). Конкретные объекты, составляющие класс, называют экземплярами соответствующей сущности. Между сущностями идентифицируются взаимосвязи различного вида: один к одному, один ко многим и др.

В силу своей ориентации на процесс проектирования ER-модели могут рассматриваться как обобщение и развитие иерархических и сетевых моделей. Это, в частности, означает, что допускаются представления связей типа "один к одному" (1:1), "один ко многим" (1 : M), "многие к одному" (M : 1) "многие ко многим" (M : N).

Цель инфологического моделирования – обеспечение наиболее естественных для человека способов сбора и представления той информации, которую предполагается хранить в создаваемой базе данных.

Инфологическая модель является промежуточной между моделью построенной на естественном языке и компьютерной реализацией модели. Модели на естественном языке не могут быть использо-



ваны в чистом виде из-за неоднозначности любого естественного языка.

По этой причине инфологическую модель данных строят по аналогии с естественным языком. Основными конструктивными элементами инфологических моделей являются сущности, связи между ними и их свойства (атрибуты).

Инфологическое моделирование включает:

- анализ объектов предметной области,
- их классификацию,
- формализованное описание сущностей и связей,
- построение структурной схемы.

Структурная схема (граф) служит основой построения базы данных.

**Сущность** – любой различимый объект (объект, который мы можем отличить от другого), информацию о котором необходимо хранить в базе данных. Сущностями могут быть объекты и явления земной поверхности, данные, характеристики объектов и т.д.

Различают тип сущности и экземпляр сущности. Например, типом сущности может быть СУБЪЕКТ ФЕДЕРАЦИИ, а экземпляром – Москва, Московская область и т.д.

Характеристикой сущности является **Атрибут**. Его наименование должно быть уникальным для конкретного типа сущности. В тоже время атрибут может быть одинаковым для различного типа сущностей. Например, атрибут ЦВЕТ может быть определен для многих сущностей: ФОТОСНИМОК, ТОПОГРАФИЧЕСКАЯ КАРТА, СПЕКТРОЗОНАЛЬНЫЙ СНИМОК, ТЕМАТИЧЕСКАЯ КАРТА, ТЕПЛОВОЙ СНИМОК и т.д.

Атрибуты используются для определения типов информации, которая должна быть собрана о сущности. Примерами атрибутов для сущности СУБЪЕКТ ФЕДЕРАЦИИ являются КАДАСТРОВЫЙ НОМЕР, КОД ЭКОНОМИЧЕСКОГО РЕГИОНА, ПЛОЩАДЬ, ЧИСЛЕННОСТЬ НАСЕЛЕНИЯ и т.д. Для атрибута также существует различие между типом и значением. Тип атрибута кадастровый номер имеет значения от 1 до 89. Однако каждому экземпляру сущности присваивается только одно значение атрибута.

В базе данных ГИС сущность отображается одной строкой таблицы. Абсолютное различие между типами сущностей и атрибутами отсутствует. Атрибут является таковым только в связи с типом сущности. В другом контексте атрибут может выступать как самостоятельная сущность.

**Ключ** – это тот минимальный набор атрибутов, по значениям которых можно однозначно найти требуемый экземпляр сущности. Минимальность означает, что исключение из набора любого атрибута не позволяет идентифицировать сущность по оставшимся. Для сущности СУБЪЕКТ ФЕДЕРАЦИИ (СФ) ключом является атрибут Кадастровый номер.

**Связь** – ассоциирование двух или более сущностей. Одно из основных требований к организации базы данных – это обеспечение возможности отыскания одних сущностей по значениям других, для чего необходимо установить между ними определенные связи. А так как в реальных базах данных нередко содержатся сотни или даже тысячи сущностей, то теоретически между ними может быть установлено более миллиона связей. Наличие такого множества связей и определяет сложность инфологических моделей.

При построении инфологических моделей можно использовать язык *ER-диаграмм*. В них сущности изображаются помеченными прямоугольниками, ассоциации – помеченными ромбами или шестиугольниками, атрибуты – помеченными овалами, а связи между ними – ребрами, над которыми может проставляться степень связи как и в иерархической модели.

Связь ОДИН-К-ОДНОМУ (1:1) описывает ситуацию, когда каждому экземпляру сущности А соответствует один представитель сущности другой сущности В. Например, если сущность А - СУБЪЕКТ ФЕДЕРАЦИИ, то ей соответствует представитель сущности В - АДМИНИСТРАТИВНЫЙ ЦЕНТР.

Связь ОДИН-КО-МНОГИМ (1:М) описывает ситуацию, при которой а сущности А соответствуют 0, 1 или несколько представителей сущности В.

Например, если сущность Г - АДМИНИСТРАТИВНЫЙ РАЙОН, то ей соответствуют представители сущности Д - РАЙОННЫЕ ПОСЕЛКИ.

Так как между двумя сущностями возможны связи в обоих направлениях, то существует еще два типа связи МНОГИЕ-К-ОДНОМУ (М:1) и МНОГИЕ-КО-МНОГИМ (М:М).

Иногда применяется менее наглядный, но более содержательный язык инфологического моделирования (ЯИМ), в котором сущности и ассоциации представляются предложениями вида:

СУЩНОСТЬ (атрибут 1, атрибут 2, ..., атрибутN)  
 АССОЦИАЦИЯ [СУЩНОСТЬ S1, СУЩНОСТЬ S2, ...]  
 (атрибут 1, атрибут 2, ..., атрибут N)

где  $S$  – степень связи, а атрибуты, входящие в ключ, должны быть отмечены с помощью подчеркивания.

К. Дейт [13] определяет три основных класса сущностей: *стержневые, ассоциативные и характеристические*, а также подкласс ассоциативных сущностей – *обозначения*.

**Стержневая сущность** – характеризуется как независимая базовая сущность

**Ассоциативная сущность (ассоциация)** – это связь вида "многие-ко-многим" между сущностями или экземплярами сущностей.

Ассоциации рассматриваются как полноправные сущности и обладают следующими свойствами:

- могут участвовать в других ассоциациях и обозначениях точно так же, как стержневые сущности;
- могут обладать свойствами, т.е. иметь не только набор ключевых атрибутов, необходимых для указания связей, но и любое число других атрибутов, характеризующих связь.

**Характеристической сущностью (характеристикой)** – называют связь вида "многие-к-одной" или "одна-к-одной" между двумя сущностями (частный случай ассоциации).

Единственная цель характеристики в рамках рассматриваемой предметной области состоит в описании или уточнении некоторой другой сущности. Необходимость в них возникает в связи с тем, что сущности реального мира имеют иногда многозначные свойства. Например, цифровая модель – несколько характеристик ее формирования (исходная, дополненная, переработанная, ... и т.д.).

Существование характеристики полностью зависит от характеризуемой сущности.

Для описания характеристики используется новое предложение ЯИМ, имеющее в общем случае вид:

**ХАРАКТЕРИСТИКА (атрибут 1, атрибут 2, ...)  
{СПИСОК ХАРАКТЕРИЗУЕМЫХ СУЩНОСТЕЙ}.**

Описание характеристики на ЯИМ строится так, что атрибуты помещают в круглые скобки, сущности в фигурные

**Обозначающая сущность (обозначение)** – это связь вида "многие-к-одной" или "одна-к-одной" между двумя сущностями и отличается от характеристики тем, что не зависит от обозначаемой сущности.

Обозначения и характеристики не являются полностью независимыми сущностями, поскольку они предполагают наличие некоторой другой сущности, которая будет "обозначаться" или "характери-

зоваться". Однако они все же представляют собой частные случаи сущности и могут иметь свойства, могут участвовать в ассоциациях, обозначениях и иметь свои собственные (более низкого уровня) характеристики.

Рассмотрим пример, связанный с зачислением сотрудников в различные отделы организации.

При отсутствии жестких правил (сотрудник может одновременно зачисляться в несколько отделов или не зачисляться ни в один отдел) создаем описание ассоциативной сущности Зачисление:

Отделы (Номер отдела, Название отдела, ...)

Служащие (Табельный номер, Фамилия, ...)

Зачисление [Отделы M, Служащие N]

(Номер отдела, Табельный номер, Дата зачисления).

Сущности помещаются в квадратные скобки, атрибуты - в круглые.

Если каждый из сотрудников должен быть обязательно зачислен в один из отделов, можно создать описание обозначающей сущности Служащие:

Отделы (Номер отдела, Название отдела, ...)

Служащие (Табельный номер, Фамилия, ... , Номер отдела, Дата зачисления)[Отделы]

В данном примере служащие имеют независимое существование (если удаляется отдел, то из этого не следует, что также должны быть удалены служащие такого отдела). Поэтому они не могут быть характеристиками отделов и названы обозначениями.

Обозначения используют для хранения повторяющихся значений больших текстовых атрибутов.

Описание обозначения внешне отличается от описания характеристики только тем, что обозначаемые сущности заключаются не в фигурные скобки, а в квадратные:

ОБОЗНАЧЕНИЕ (атрибут 1, атрибут 2, ...)

[СПИСОК ОБОЗНАЧАЕМЫХ СУЩНОСТЕЙ].

Как правило, обозначения не рассматриваются как полноправные сущности.

Для выявления связей между сущностями необходимо, как минимум, определить сами сущности. Но это не простая задача, так как в разных предметных областях один и тот же объект может быть сущностью, атрибутом или ассоциацией.

Что же такое "связь"? В ER-диаграммах это линия, соединяющая геометрические фигуры, изображающие сущности, атрибуты,

ассоциации и другие информационные объекты. В тексте же этот термин используется для указания на взаимозависимость сущностей. Если эта взаимозависимость имеет атрибуты, то она называется *ассоциацией*.

### 3.5. Сетевые модели

Сетевые модели дают представление об объектах, связанных отношениями "многие ко многим". В отличие от иерархических моделей в сетевой каждый объект может иметь несколько "подчиненных" и несколько "старших" объектов.

Сетевые модели используют табличные и значительно чаще графовые представления. Вершинам графа сопоставляют некоторые типы сущности, представляемые таблицами, а дугам - связи.

Многие типы сетевых моделей данных используют для описания экономических и организационных систем.

Рассмотрим на примере построение сетевой модели. При использовании этих моделей для решения задач планирования узлы называют состояниями, так как они соответствуют состояниям системы или состояниям процесса выполнения работ. Дуги называют работами, так как они соответствуют определенным этапам работ при переходе от одного состояния к другому.

Дуги ориентированы, т.е. показывают направление. Обычно длина дуги соответствует численному значению длительности данного этапа работы.

Табличные данные обычно являются исходными для построения графовой модели и их называют условиями.

На рис. 3.6. приведены табличные условия, требующие реализации в виде графа. Для построения сетевой модели необходимо определить возможные состояния системы и возможные виды работ. Дуги и узлы обозначают по разному, например, узлы (состояния) цифрами, а дуги (работы) буквами.

Рассмотрим практический пример составления сетевой модели для описания процесса решения задачи при помощи компьютера и передачи результатов заказчику.

Таблица, приведенная на рис. 3.6, содержит необходимые виды работ и состояния. В этой же таблице введены обозначения. Цифрами обозначены (идентифицированы) состояния, а буквами - работы.

№	Состояния	ID	Работы
1	Начало работы	А	Формулировка задачи
2	Задача сформулирована	Б	Поиск решений
3	Завершение поиска	В	Выбор модели
4	Модель найдена	Г	Построение модели
5	Модель построена	Д	Сбор данных для расчета
6	Сбор данных завершен	Е	Создание программы
7	Программа составлена	Ж	Расчеты
8	Расчеты завершены	И	Передача результатов заказчику
9	Окончание работы		

Рис. 3.6. Исходные данные для построения сетевой модели

На основе приведенных данных строится сетевая модель. Эта модель приведена на рис.3.8. Она читается следующим образом. На начальном этапе 1 осуществляется формулировка задачи. Предполагается, что задачу можно сформулировать и это соответствует состоянию 2. После этого осуществляется поиск готового решения, выполняется работа Б. После завершения поиска состояние 3 возможны два результата: решение найдено, готовое решение не найдено.

Наличие двух возможных вариантов на схеме отображается двумя путями, т.е. дугами. Если решение найдено, то происходит переход к состоянию 6. Если решение не найдено, то выполняют работу В по поиску моделей которую можно использовать для решения задачи.

В результате выполнения этой работы также возможны два варианта, что приводит к ветвлению, т.е. к появлению двух дуг из узла, соответствующего состоянию 4.

Если модель найдена, осуществляют сбор информации для получения решения и переходят к состоянию 6. Если модель не найдена, осуществляют ее построение, переходят в состояние 5 и после этого осуществляют сбор информации для получения решения и переходят к состоянию 6.

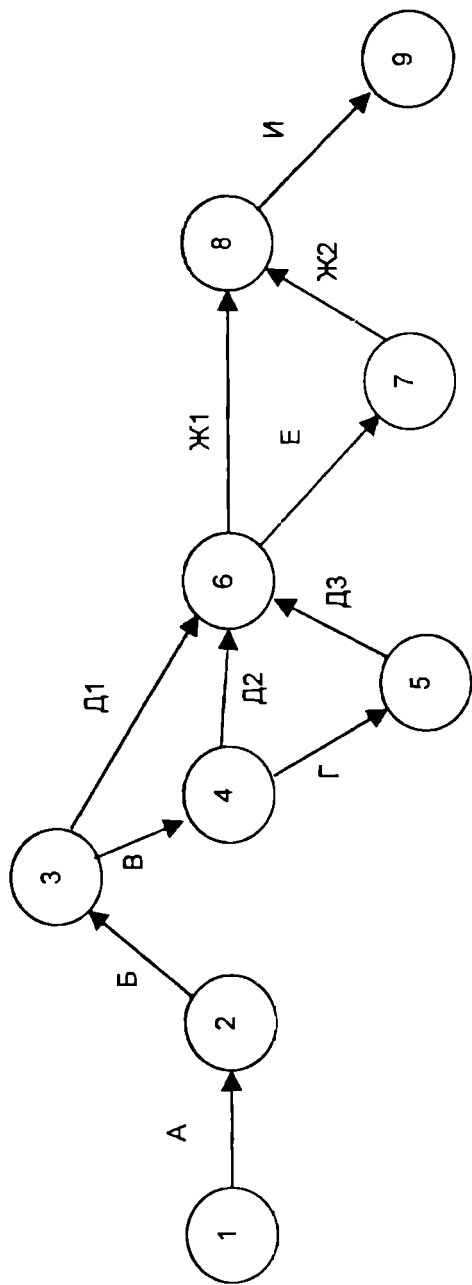


Рис. 3.7. Графовая схема сетевой модели

Следует обратить внимание, что работа по сбору информации, обозначенная в таблице рис.3.7 идентификатором Д, на практике имеет три варианта Д1, Д2, Д3. Это означает, что объем данной работы для разных случаев различен.

В состоянии 6 возможны также два варианта: программа для вычисления готова, программа для вычисления не готова. В первом случае осуществляют расчет (работа Ж1) и переходят в состояние 8. Во втором случае составляют программу (работа Е) и переходят в состояние 7.

Из состояния 7 переходят в состояние 8 и из него в конечное состояние 9.

Таким образом, в сетевой модели возможно ветвление как в иерархической и объединение, что в иерархической модели недопустимо.

Ветвление в сетевой модели возможно в двух случаях:

- возникновение альтернативного варианта действий;
- возможность параллельных (независимых) действий.

Объединение дуг в сетевой модели также чаще всего возникает в двух вариантах:

- объединение результатов действия альтернатив (разные альтернативы приводят к одному результату);
- окончание параллельных независимых действий (результаты независимых действий суммируются).

Следует отметить, что сетевая модель используется и при описании пространственных объектов. Примером сетевой модели может служить сеть Интернет.

Основным назначением иерархической модели является выделение функциональных уровней иерархии моделируемого и определение подчиненности узлов (элементов) одного уровня другому.

Основным назначением модели "сущность - связь" является выделение элементов моделируемого объекта, которые относят к "сущностям" и выделение и определение связей между этими сущностями.

Сетевые модели применяют в двух случаях. Одним из назначений сетевых моделей является выделение состояний, в которых может находиться моделируемый объект, и действий (работ) необходимых для перехода из одного состояния в другое.

Другим назначением сетевых моделей является отражение связей и отношений между пространственными объектами. Именно для сетевых моделей, описывающих пространственные отношения, важной характеристикой становится топология, которая будет рассмотрена позже.

Таким образом, рассмотренные модели дополняют друг друга и в совокупности дают более полное описание моделируемого объекта.



Анализ структурных базовых моделей в геоинформатике дает основание сделать вывод о том, что аналогичные модели используются в информационных системах, решающих экономические задачи и задачи управления.

Следовательно, на уровне базовых моделей ГИС (как и ОАСУ) применимы для решения разнообразных задач управления объектами.

#### 4. Сбор информации

Обязательным компонентом любой системы обработки и хранения данных является входной поток первичной информации, поступающей из внешней среды. Анализ внешней информационной среды предполагает: выявление источников необходимой информации и связей этих источников с информационной системой потребителя, именно:

- оценку надежности источников информации;
- оценку достоверности информации, которой обладает источник;
- определение объемов и формы представления информации у источника;
- выяснение условий и особенностей предоставления информации источником;

Источниками информации могут служить любые объекты окружающего нас мира: предметы живой и неживой природы, события, явления, процессы (социальные, экономические и проч.), а также сами информационные системы и технологии. Источниками информации могут быть субъекты, владеющие сведениями или профессиональными знаниями в определенной предметной области.

Выбор источников информации и технологий ее сбора определяется результатами анализа внешней информационной среды и техническими средствами, имеющимися в распоряжении исследователя. В большинстве случаев приходится использовать несколько информационных источников, дополняющих друг друга. При альтернативных источниках информации критериями их выбора служат следующие характеристики:

- надежность и доступность источника;
- необходимость и достоверность предоставляемой информации;
- стоимость информационных услуг;
- совместимость формы существующего у источника представления информации с требованиями информационной системы потребителя.

Получение информации от источника осуществляется, как правило, на основе каких-либо затрат. Размер затрат определяется типом тех-

нологии. Поэтому оптимальный выбор технологий сбора информации при разных информационных источниках снижает затраты.

При сборе информация регистрируется, кодируется, аннотируется, систематизируется и унифицируется.

Среди методов сбора неунифицированной информации можно выделить:

- непрерывный мониторинг процессов и явлений;
- статистический анализ информации об объектах;
- получение информации по системам телекоммуникаций;
- электронный поиск в информационных сетях;
- информационная разведка и др.

При мониторинге наблюдаемые данные должны фиксироваться в удобной для контроля и последующей обработки форме.

Целью статистического анализа, как правило, является получение оценок параметров или характеристик внешней или внутренней среды исследуемой системы.

Прием и регистрация собираемой первичной информации на входе информационной системы могут производиться вручную, автоматизированным способом или автоматически. При ручном способе данные фиксируются на специальных бланках и в журналах в установленной табличной форме, а поступающие документы регистрируются и сохраняются в виде оригиналов или копий. При автоматизированном способе регистрация осуществляется на машинном носителе с помощью автоматизированных или полуавтоматизированных устройств путем диалога.

Автоматизированная регистрация выполняется без участия человека и предполагает прямое подключение информационной системы к источнику. В этом случае входная информация представляет собой кодированный сигнал.

При любом способе регистрации входная информация должна контролироваться с целью недопущения ее дублирования или обнаружения в ней ошибочных данных.

Собранная таким образом информация представляет собой первичные данные. Затем эти данные подлежат дальнейшей формализации и унификации.

При проектировании конкретной информационной системы или технологии проблема выбора способа сбора информации при наличии альтернатив должна решаться в контексте общей проблемы технико-экономического обоснования. Это связано с возможной зависимостью затрат на последующую обработку информации от формы ее представления, диктуемой избранным способом сбора.

## 4.1. Первичная обработка информации

Накопленная первичная информация должна пройти предварительную (первичную) обработку для возможности ее комплексного использования, устранения избыточности, систематизации и хранения в базе данных. Другими словами, из разрозненных наборов данных должен быть сформирован единый набор или информационная основа.

Комплексная информация позволяет пользователю любой информационной системы принимать более обоснованные решения. Такой комплексной информацией могут служить совокупности экономических, технических и др. показателей, параметры математических моделей, значения факторов, бездействующих на управляемый социально-экономический, технический или иной процесс, геодезические или географические координаты и многое другое.

В зависимости от способа получения, достоверности и объема исходной информации при ее обработке могут применяться как методы статистического анализа, так и математические методы обработки детерминированных данных.

Среди статистических методов обработки можно выделить две группы методов: предварительной обработки и статистического анализа. К первым методам относят группировку и сводку данных. Группировка данных представляет собой выделение части данных из общей совокупности по одному или сочетанию нескольких признаков. Сводка (суммирование) данных обеспечивает их количественную оценку как по всей совокупности этих данных (простая сводка), так и по отдельным группам этой совокупности (групповая сводка).

Методы статистического анализа позволяют получать оценки более "скрытых" характеристик процессов и явлений.

В отличие от статистических методов математические методы обработки детерминированной информации применяются для выявления параметров заданных функций.

Программное обеспечение, реализующее эти методы обработки, может создаваться под конкретную информационную систему или в виде прикладных программных пакетов (ППП), не зависящих от содержания предметной области и предназначенных для использования специалистами, способными формализовать решаемые задачи на языке представления данных.

## 4.2. Кодирование информации

В информационных технологиях осуществляется обрабатывать числовой, текстовой, графической, аудио и видео информации.

Все эти виды информации кодируются в последовательности нулей и единиц. Такое кодирование информации называется двоичным кодированием, а логические последовательности нулей и единиц образуют машинночитаемый язык.

Двоичный код имеет алфавит, который состоит из двух цифр (0 и 1). Эти цифры можно рассматривать как два состояния (события). При записи двоичной цифры реализуется выбор одного из двух возможных состояний и, следовательно, она несет количество информации равное 1 бит. Каждая цифра машинного двоичного кода несет количество информации равное одному биту.

В процессе обработки информации в компьютерных технологиях широкое распространение получили также восьмеричная и шестнадцатеричная системы счисления. Программисты часто используют эти системы представления на этапах отладки программ или просмотра содержимого файлов, для того, чтобы информация представить информацию в более удобной для просмотра сжатой форме (используется меньшее количество разрядов).

Числа могут быть записаны в *естественной, целой или нормализованной* формах. Естественной формой называется запись чисел, включающая все разряды и разделитель дробной и целой части, например, 3.14 или 10000. Этот вид записи удобен для чтения, но практически не используется при расчетах.

Целая форма основана на исключении дробной части при расчетах и кодировании чисел. Она позволяет проводить более скоростные операции вычислений.

Нормализованная форма чисел обычно используется для записи чисел, для которых необходимо сохранение дробной части. Ее удобно использовать в тех случаях, когда в обычной естественной форме числа содержат большое количество незначащих нулей (например,  $0,0001 = 0,1 \cdot 10^{-3}$  или  $10000 = 10^4$ ).

Нормализованная форма использует экспоненциальную запись чисел, при которой указывается мантисса, основание системы счисления и степень. Число  $A$  в любой системе счисления в экспоненциальной форме имеет вид:

$$A = m \cdot q^n,$$

где  $m$  — мантисса числа, правильная дробь, имеющая после запятой цифру отличную от нуля,  $q$  — основание системы счисления,  $n$  — порядок числа.

Нормализованная форма обеспечивает большую точность вычислений, но меньшую скорость. При обработке чисел в целой форме диапазон их изменения меньше чем при нормализованной форме.

Для кодирования одного символа текстовой информации используется количество информации равное 1 байту (8 бит).

В одном байте можно закодировать 256 различных символов. Такого количества символов вполне достаточно для представления текстовой информации, включая прописные и заглавные буквы русского и латинского алфавита, цифры, знаки, графические символы и т.д.

Кодирование текста заключается в том, что каждому символу ставится в соответствие уникальный десятичный код от 0 до 255 или соответствующий ему двоичный код от 00000000 до 11111111. Таким образом, различия в начертании символов преобразуются в различия между их кодом. В сущности этот же принцип заложен при кодировании графической информации.

При вводе в память компьютера текстовой информации происходит ее двоичное кодирование, т.е. символ преобразуется в двоичный код. При выводе символа на экран компьютера производится обратный процесс — декодирование, что в данном случае означает преобразование двоичного кода символа в его изображение.

При вводе в компьютер графической или звуковой информации она превращается в совокупность дискретных значений - квантуется. Затем эти дискретные значения кодируются подобно символьным последовательностям.

Присвоение символу или дискретному значению графического образа конкретного кода — это вопрос соглашения, которое фиксируется в кодовой таблице. Если для текста существует набор стандартов кодирования, то при кодировании графики один и тот же образ может быть закодирован по-разному в зависимости от используемого программного обеспечения. Таким образом получаются разные форматы графических данных.

При кодировании символов принято, что первые 33 кода (с 0 по 32) соответствуют не символам, а операциям (перевод строки, ввод пробела и т.д.).

Коды с 33 по 127 являются интернациональными и соответствуют символам латинского алфавита, цифрам, знакам арифметических операций и знакам препинания.

Коды с 128 по 255 являются национальными, т.е. в национальных кодировках одному и тому же коду соответствуют различные символы. К сожалению, в настоящее время существуют пять различных кодовых таблиц для русских букв (КОИ-8, CP1251, CP866, Mac, ISO), поэтому тексты, созданные в одной кодировке, не будут правильно отображаться в другой.

Каждая кодировка задается своей собственной кодовой таблицей. Одному и тому же двоичному коду в различных кодировках поставлены

в соответствии различные символы. Эти кодировки являются аналогами форматов графических данных. Кодировки играют важную роль в текстовых процессорах и в системах электронной почты.

Необходимо остановиться более подробно на кодировании графической информации, поскольку в геоинформатике она играет важную роль. Кодирование изображения основано на его пространственной дискретизации, в то время как кодировании аудиосигнала – на временной.

Пространственная дискретизация изображения основана на преобразовании изображения в мозаику (цветную или черно-белую). Изображение разбивается на отдельные маленькие элементы (пиксели), при этом каждому элементу присваивается значение его плотности (код плотности) и цвета (код цвета красный, зеленый, синий).

Качество кодирования изображения зависит от двух параметров. Во-первых, оно тем выше, чем меньше размер пикселя, что существенно увеличивает объем памяти для хранения такого изображения.

Во-вторых, чем большее количество цветов, тем более качественно кодируется изображение (каждая точка несет большее количество информации). Совокупность используемого набора цветов образует *палитру цветов*.

Таким образом, при вводе любого изображения в компьютер возникает необходимость его дискретизации и кодирования. При этом, как и для символов, следует выполнить обратный процесс декодирования, т.е. вывести изображения на экран.

Для решения этой задачи графическая информация на экране монитора формируется в виде набора строк (растра). Такое изображение называется растровым изображением и в современных информационных технологиях оно формируется уже на стадии ввода при помощи сканеров. Растровое изображение создается из определенного количества строк, которые, в свою очередь, содержат определенное количество точек (пикселей).

Качество вывода изображения определяется разрешающей способностью монитора, т.е. количеством строк на дюйм. Чем больше разрешающая способность, т.е. чем больше количество строк растра и точек в строке, тем выше качество изображения. В современных персональных компьютерах обычно используются четыре основных разрешающих способности экрана: 640 x 480, 800 x 600, 1024 x 768 и 1280 x 1024 точки.

Цветные изображения формируются в соответствии с двоичным кодом цвета каждой точки, хранящимся в видеопамати. Цветные изображения могут иметь различную *глубину цвета*, которая задается используемым количеством бит для кодирования цвета точки. Наиболее

распространенными значениями глубины цвета являются 4, 8, 16 или 24 бита на точку.

С начала 90-х годов в информационных технологиях широко стала применяться работа со звуковой (аудио) информацией. Каждый компьютер, имеющий звуковую плату, микрофон и колонки, может записывать, сохранять и воспроизводить звуковую информацию.

С помощью специальных программных средств (редакторов аудиофайлов) открываются широкие возможности по созданию, редактированию и прослушиванию звуковых файлов. Создаются программы распознавания речи, в которых появляется возможность преобразования речи в набор кодов и последующее преобразование кодов в символы текста для печати и хранения.

Звук представляет собой волну с непрерывно меняющейся амплитудой и частотой. Амплитуда сигнала определяет громкость звука, частота сигнала определяет высоту тона. Для обработки непрерывный звуковой сигнал, как и символы и изображения, должен быть закодирован, то есть превращен в последовательность двоичных нулей и единиц.

При кодировании непрерывного звукового сигнала производится его дискретизация по времени. Непрерывная звуковая волна разбивается на отдельные маленькие временные участки, причем для каждого такого участка устанавливается определенная величина амплитуды. Этот процесс можно сравнить с преобразованием гладкой кривой линии в гистограмму.

При этом непрерывная зависимость амплитуды сигнала от времени  $A(t)$  заменяется на дискретную последовательность уровней громкости. На графике это выглядит как замена гладкой кривой на последовательность «уровней». В теории моделирования этот процесс называют заменой динамической модели на квазистационарную.

Уровни соответствуют громкости звука и кодируются подобно кодированию полутонового изображения. Уровни звука можно рассматривать как набор возможных состояний, соответственно, чем большее количество уровней громкости будет выделено в процессе кодирования, тем более точная передача по громкости звукового сигнала будет при его воспроизведении.

Современные звуковые карты обеспечивают 16-битную глубину кодирования звука. Количество различных уровней сигнала или состояний при данном кодировании равно  $N = 2^{16} = 65536$ .

Таким образом, современные звуковые карты могут обеспечить кодирование 65536 уровней сигнала. Каждому значению амплитуды звукового сигнала присваивается 16-битный код.

Поскольку при двоичном кодировании непрерывный сигнал заменяется на последовательность дискретных уровней, то качество коди-

вания определяется количеством квантований сигнала в единицу времени, т.е. частотой дискретизации. Чем больше частота дискретизации, тем более информативно передается звуковой сигнал и тем качественнее его воспроизведение по частотным характеристикам.

Частота дискретизации аналогового звукового сигнала может принимать значения от 8 до 48 кГц. При частоте 8 кГц качество дискретизированного звукового сигнала соответствует качеству радиотрансляции, а при частоте 48 кГц — качеству звучания аудио-CD.

## 5. Интеграция данных в геоинформатике

Модели могут иметь разные цели создания, разные формы представления и разный принцип организации. Для обработки в единой технологической системе или в единой информационной системе данные должны быть объединены на основе какого-либо правила или метода. Таким объединяющим методом является интеграция данных [17, 21, 22].

Прежде чем дать определение интеграции необходимо отметить, что в геоинформатике совокупность изучаемых и исследуемых параметров данных и моделей данных образует систему. Это обусловлено тем, что эта информационная совокупность отображают реальные объекты и явления земной поверхности, которые расположены не произвольно, а организовано. Это дает основание строить информационно-ресурсную модель.

Можно говорить, что информационно-ресурсная модель об объектах и явлениях земной поверхности образует некую систему. Отдельные модели или данные являются элементами такой системы.

Интеграцией в системе называют восстановление и/или повышение качественного уровня взаимосвязей между элементами системы, а также процесс создания из нескольких разнородных систем единой системы, с целью исключения (до технически необходимого минимума) функциональной и структурной избыточности и повышения общей эффективности функционирования [11].

Интегрированная модель является разновидностью информационно-ресурсной модели с высокой степенью ее организованности. Она не является просто суммой информационных частей ее образующих. Как правило, она имеет меньший объем физической памяти при сохранении информационной емкости по сравнению с информационными моделями, ее составляющими, хотя включает данные о связях и дополнительную служебную информацию.

Кроме того, интегрированная модель не только описывает информационные свойства объекта, но и позволяет проводить эффективную обработку данных, относящихся к исследуемому объекту. В этом за-



ключается преимущество интегрированной модели по сравнению с другими.

Упрощенно интеграцию можно представить в виде следующих процедур:

Выявление или классификацию элементов неоднородной системы;

Выбора или создание интегрированной информационной основы;

Объединение элементов с исключением информационной избыточности

Преобразование элементов системы к единому виду;

Установление связей между элементами системы.

Интегрированная модель является частным случаем информационно-ресурсной модели. Отличие интегрированной модели от других информационно-ресурсных моделей в том, что она имеет некую основообразующую переменную, которая служит средством связи между всеми остальными переменными. Эта переменная должна обладать наибольшей устойчивостью по сравнению с остальными.

Следовательно, прежде чем осуществить интеграцию необходимо выделить основные группы данных и выбрать среди них те, которые могут послужить основой для объединения (интеграции) всех остальных.

Проведем анализ данных для выбора основообразующей переменной. Данные, используемые в геоинформатике, рассматривают с учетом трех аспектов: пространственного, временного и тематического.

Пространственный аспект связан с определением местоположения, временной - с изменениями объекта или процесса с течением времени, в частности от одного временного среза до другого. Тематический аспект обусловлен выделением одних признаков объекта и исключением из рассмотрения других.

Данные, которые выбирают для интеграции, должны быть наиболее устойчивыми или наименее изменяющимися. Временные данные по определению изменяющиеся и поэтому не могут служить основой интеграции.

Тематические данные также могут меняться, исчезать или появляться в новых видах, поэтому и они не могут служить основой интеграции.

Следовательно, пространственные данные - наиболее устойчивые и наименее меняющиеся, поэтому в этой группе следует искать основу для интеграции. Среди пространственных данных наиболее устойчивыми (наименее изменчивыми) являются географические координаты. Именно они являются основой для объединения различных данных, т.е. основой для интеграции.

Следует отметить, что интеграция данных в геоинформатике более общее понятие, чем интеграция данных в информационной системе, например в ГИС. Это обусловлено тем, что интеграция данных в геоинформатике включает все изучаемые данные, а в ГИС только данные, используемые при обработке и хранении в этой информационной системе.

Кроме того, интеграция данных в геоинформатике решается на георетическом уровне, а в ГИС на технологическом, что накладывает ограничения на возможность реализации всех интеграционных связей.

## 6. Теоретические основы обработки данных в геоинформатике

### 6.1. Информационные процессы

Процесс в общем случае состоит из множества событий, которые соответствуют выполнению действий, и возникают в определенном причинно-следственном (временном) порядке. Процесс или совокупность процессов описывает возможное функционирование сложной системы.

В геоинформатике применяют различные системы, которые могут быть геоинформационными или просто информационными. Эти системы могут быть локализованными или распределенными. Информационный процесс описывает возможное функционирование (распределенной) информационной или геоинформационной системы.

Информационный процесс служит для математического моделирования любого дискретного действия геоинформационной системы и складывается из отдельных действий. Между этими действиями существуют причинно-следственные отношения. В общем случае система может быть распределенной, поэтому все дальнейшие рассуждения будем относить к распределенным системам, имея в виду, что локальные системы можно рассматривать как частный случай.

При каждой реализации действия говорят о событии. Действие распределенной системы состоит из множества событий, причем каждому событию предписано действие. Различным событиям могут быть предписаны одни и те же действия. Это приводит к следующей математической модели (структуре) для процесса.

Пусть задано множество (универсум)  $E$  событий (events), множество  $A$  действий (actions). Тройка  $p=(E_0, \leq_0, \alpha)$  называется структурой действия или также процессом, если справедливы следующие высказывания:

$$E_0 \subseteq E,$$

$\leq_0$  есть частичный порядок над  $E_0$ ,

$$\alpha: E_0 \rightarrow A.$$

Здесь  $E_0$  называется множеством событий,

$\leq_0$  есть отношение причина-следствие,

$\alpha$  - обозначение действия процесса  $p$ .

Отображение  $\alpha$  каждому событию предписывает действие. В

это определение структуры действия включается также пустой процесс. Он представляется структурой действия, которая содержит пустое множество событий.

В случае сложных конечных или бесконечных процессов рассмотрение неструктурированного множества всех событий и связанных с ними действий становится сложным. Часто процесс складывается из ряда частичных процессов (подпроцессов). Такого рода процесс можно анализировать путем разложения на подпроцессы и их анализа.

Процессы могут состоять из совокупности последовательных (зависимых) событий или из совокупности независимых событий. В первом случае процессы называют *последовательными*, во втором - *параллельными*. Последовательные процессы являются частным случаем параллельных процессов.

Последовательные процессы могут быть описаны с помощью более простой формы структуры действий. Представления последовательных процессов описывают как конечную или рекурсивную последовательность действий.

Каждому последовательному процессу однозначно сопоставим поток действий (spug) – трасса (trace) процесса. Трассы являются более простыми моделями для последовательного хода работы системы, чем структуры действий.

Для моделирования распределенных систем, как правило, используются не процессы с их явным представлением параллельности (true concurrency – истинный параллелизм), а технически более простые для применения множества потоков действий. В этом случае говорят об искусственной линейаризации (interleaving).

Абстрактное описание динамического поведения распределенной системы может быть дано путем задания множества процессов. Существует ряд способов для описания систем и ходов их работ:

- сети Петри, графический метод описания;
- агенты, формальный язык описания;
- формулы логики предикатов для описания хода работы.

Одной из первых концепций, которые были предложены для описания распределенных систем и процессов, являются сети Петри, представляющие собой простое графическое представление таких систем.

**Сеть Петри (сеть условий/событий)** – направленный граф, состоящий из узлов двух типов – вентилях (переходы, барьеры) и ячеек

(места, площадки). Ребра по мере надобности ведут от вентилях к ячейкам или от ячеек к вентилям. Ячейки загружаются логическими значениями или натуральными числами. Содержимое ячеек определяет состояние сети. В каком-либо заданном состоянии определенные вентиля (множество вентилях) готовы к передаче (они могут «открываться» - путем открытия такого множества вентилях изменяя загрузку сети, т.е. содержимое ее ячеек).

Для изучения информационных процессов необходимо ввести ряд основных понятий, относящихся к моделированию, описанию и программированию параллельно работающих распределенных систем. Дно из таких понятий – «вычислительная система».

Вычислительная система – определенная архитектура взаимодействующих компонентов, отделенных от окружения. Если активности компонент могут происходить одновременно – то это параллельно работающие системы или параллельно протекающие (параллельные) процессы. Тех случаях, когда системы построены из отдельных, удаленных друг от друга в пространстве компонент, тогда речь идет об распределенных системах, или об распределенных процессах.

Для описания процессов применяются следующие термины:

- **распределение** означает пространственное расположение (или разделение) отдельных компонент процесса;
- **параллелизм** относится к временным отношениям между действиями компонент процесса, которые могут протекать одновременно (параллельно);
- **интерактивность, реакция, коммуникация, координация, синхронизация** касаются причинно-следственных отношений, возникающих между пространственно разделенными и выполняемыми одновременно друг с другом действиями (с помощью них осуществляется обмен сообщениями и сигналами);
- **недетерминированность** при моделировании распределенных процессов возникает из-за целенаправленного опускания из рассмотрения определенных деталей информации, например, временных взаимозависимостей в функционировании процесса, которые являются существенными при принятии решений в ходе работы процесса.

## 6.2. Описания систем через множество процессов.

Распределенные системы обладают многими вариантами работы. Для описания функционирования распределенной системы можно использовать процесс.

Другие методы описания для распределенных систем и их поведения дают языки программирования и программы. Если программа выполняется на ЭВМ, то протекает некоторый процесс, который складывается из множества событий, соответствующих действиям при выполнении программы. Программа также описывает операционным способом процесс, причем последовательные программы описывают последовательные процессы. Для описания общих, не последовательных процессов языковых средств обычных, последовательных языков программирования оказывается недостаточно. Чтобы разрешить проблему, возникающую при описании параллельно выполняющихся систем программ (например для параллельной композиции, коммуникации между частями программы и синхронизируемого доступа к общей памяти), применяют дополнительные языковые средства.

### 6.2.1. Сети Петри

Для графического описания распределенных систем и процессов широко применяют сети Петри.

Определение сети Петри было дано выше. Рассмотрим практический пример (сеть условия-события) их использования.

На рис. 6.1 приведена сеть Петри. Прямоугольники обозначают вентили, круги - ячейки. Ячейки с меткой в виде заштрихованной области внутри обозначают в булевой сети ячейки, в которые загружено true, или 1; в остальные ячейки пусть загружено false, или 0. В сетях с ячейками, загружаемыми натуральными числами, в одной и той же ячейке может находиться несколько меток.

В заданной на рис. 6.1 сети готовы к открытию вентили а и в, так как все ячейки на их входящих ребрах загружены положительно.

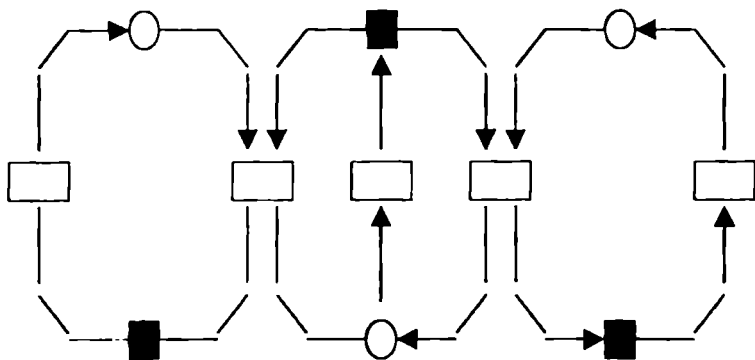


Рис. 6.1. Сеть Петри для задачи производитель-потребитель.

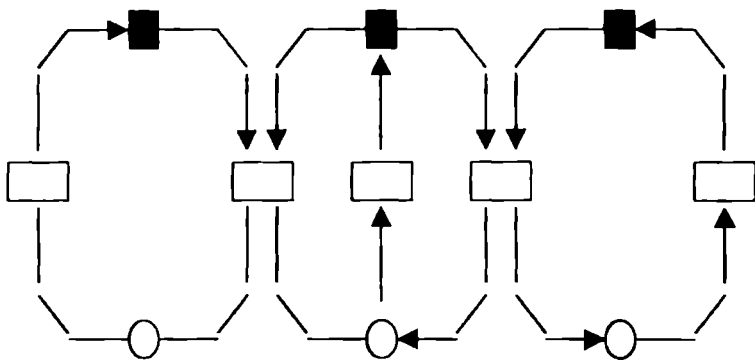


Рис. 6.2. Сеть Петри для задачи производитель-потребитель с конфликтной загрузкой.

После того как оба эти вентиля будут открыты, получим сеть, показанную на рис. 6.2, с указанной там загрузкой. Как видим, к передаче вновь готовы определенные вентили (на этот раз  $c$  и  $d$ ). Может открыться вентиль  $c$  или  $d$ , но не оба сразу.

Пусть дано множество  $T$  вентиляей и множество  $P$  ячеек ( $P \cap T = \emptyset$ ). Сеть Петри есть тройка  $(T_0, P_0, R)$ , такая, что справедливо

$$\begin{array}{ll} T_0 \subseteq T & \text{вентили,} \\ P_0 \subseteq P & \text{ячейки,} \\ R \subseteq (T_0 \times P_0) \cup (P_0 \times T_0) & \text{отношения} \end{array}$$

потока.

Множества  $T_0$  и  $P_0$  конечны. В таком случае сеть Петри есть бичастичный направленный граф. Граф называется бичастичным, если множество его узлов разбито на два класса и ребра могут соединять только узлы из различных классов.

Для сети Петри  $(T_0, P_0, R)$  каждое отображение

$$b: P_0 \rightarrow \mathbb{N},$$

которое ячейкам предписывает числа, означает целочисленную конкретизацию (занесение в ячейки конкретных чисел). Сети с целочисленными конкретизациями называются также сетями *ячейки-вентили*. Часто для каких-либо ячеек устанавливаются дополнительно определенные наибольшие значения («мощность ячейки»), которые могут иметь заносимые в эти ячейки значения.

В некоторых случаях можно ограничиться сетями Петри с конкретизациями, использующими только значения 0 и 1, или соответственно **false** и **true**. При этом рассматриваются только конкретизации вида

$$b: P_0 \rightarrow \mathbb{B}.$$

В таких случаях говорят о булевских сетях Петри или сетях условия-события.

Динамическое поведение сети может быть описано через процессы переключений. Вентили переключаются благодаря тому, что они извлекают метки из всех входных ячеек и добавляют метки во все выходные ячейки. Процесс переключения может, в частности, изменить все метки в сети. Отдельные вентили готовы к передаче, если все их входные ячейки конкретизированы строго положительно.

Для целочисленной конкретизации  $b$  сети Петри  $(T_0, p_0, R)$  непустое подмножество  $K \subseteq T_0$  называется готовым к передаче, если для каждой ячейки  $p \in P_0$  справедливо:

$$|\{k \in K: (p, k) \in R\}| \leq b(p).$$



Множество  $K$  вентиляй по этому определению для некоторой конкретизации готово к передаче, если в каждой ячейке  $p$  находится достаточно меток, чтобы все вентиля  $k \in K$ , для которых стрелка от ячейки  $p$  ведет к вентилю  $k$ , при переключении были снабжены метками. Если для каких-то ячеек предусмотрены ограничения на помещаемые туда метки, то имеет место ограничение мощности. Множество  $K$  только тогда готово к передаче, когда проходящие через вентиля метки не превосходят соответствующие максимально допустимые значения. В случае булевских сетей Петри для каждой ячейки  $p \in P_0$  и каждого вентиля  $k \in K$  требуется выполнение следующего условия для готовности к передаче множества вентиляй  $K \subseteq T$ :

$$(p, k) \in R \Rightarrow b(p) \wedge |\{k \in K : (p, k) \in R\}| \leq 1$$

и дополнительно следующего условия бесконфликтности: для каждой ячейки  $p \in P_0$  и каждого вентиля  $k \in K$  должно иметь место

$$(k, p) \in R \Rightarrow \neg b(p) \wedge |\{k \in K : (k, p) \in R\}| \leq 1.$$

Это условие говорит о том, что вентиль  $k$  может открыться только тогда, когда все ячейки, к которым ведут ребра из  $k$ , помечены через false. Аналогично булевские сети Петри соответствуют целочисленным сетям Петри, в которых каждая ячейка имеет мощность 1.

В булевских сетях Петри дополнительно требуется выполнения следующего условия «отсутствия петель»: никакой вентиль не имеет права использовать одну и ту же ячейку как в качестве входной, так и в качестве выходной ячейки. При этом должно выполняться следующее условие:

$$\forall t \in T_0, p \in P_0: \neg ((t, p) \in R \wedge (p, t) \in R).$$

Если бы для какого-либо вентиля  $t$  это условие не выполнялось, то такой вентиль (при учете правила переключения для сети условие-событие) никогда не был бы готов к передаче.

**Лемма.** Пусть заданы сеть Петри и множество  $K$  вентиляй со следующим дизъюнктивным разложением:

$$K = K_1 \cup K_2, \quad K_1 \cap K_2 = \emptyset.$$

Тогда имеет место: если множество  $K$  готово к передаче для конкретизации  $b_0$  с последующей конкретизацией  $b$ , то имеется такая конкретизация  $b_1$ , что справедливо, так как  $K_1$  готово к передаче с последующей конкретизацией  $b_1$ , и  $K_2$  готово к передаче для  $b_1$  с последующей конкретизацией  $b$ .

**Доказательство.** Определение готовности к передаче  $K$  влечет

за собой готовность к передаче любого его подмножества  $K_1$ . Для последующей конкретизации  $b_1$  при  $K_1$  для  $b_0$  подмножество  $K_2$  готово к передаче, так как  $K = K_1 \cup K_2$  было готово к передаче для  $b_0$ . Поскольку  $K_1 \cap K_2 = \emptyset$ , то сумма эффектов от  $K_1$  и  $K_2$  есть эффект  $K$ .

Сеть Петри обладает, как правило, большим числом процессов как ходов работы. Исходя из начальной конкретизации сеть Петри порождает ход работ (процесс пробегает по сети Петри) благодаря тому, что активизируются готовые к передаче множества вентилей и выполняются соответствующие им действия, в результате чего и осуществляется переход к последующим конкретизациям.

Для сети  $N = (T_0, P_0, R)$  рассмотрим процесс  $p = (E_0, \leq_0, \alpha)$  с маркировкой события через действия:

$$\alpha: E_0 \rightarrow T_0.$$

Для исходной конкретизации  $b_0$  и сети Петри  $N_0 = (T_0, P_0, R)$  конечная структура действий  $p = (E_0, \leq_0, \alpha)$  называется *ходом работы сети Петри с конечной конкретизацией*  $b_1$ , если  $p$  соответствует поведению сети с исходной конкретизацией  $b_0$ :

$$b_0 \xrightarrow{p} b_1.$$

Это отношение между конкретизацией и процессами для заданной сети определяется следующим образом:

(1) Пустой процесс всегда есть ход работы и не меняет конкретизации. Т.о. для

пустого процесса  $p$  и любой конкретизации  $b$  справедливо следующее

соотношение:

$$b \xrightarrow{p} b.$$

(2) Пусть  $p = (E_0, \leq_0, \alpha)$  - процесс с тривиальным причинным порядком, в

котором различные события являются причинно независимыми. Для

процесса  $p$  справедливо следующее:

$$\forall e, d \quad E_0: e \leq_0 d \Leftrightarrow e = d.$$

Тогда для конкретизации  $b_0$  и  $b_1$  отношение

$$b_0 \xrightarrow{p} b_1$$

справедливо только тогда, когда все события в  $p$  помечены по-разному и

потому справедлива следующая формула:

$$\forall e, d \quad E_0: e \neq d \Rightarrow \alpha(e) \neq \alpha(d),$$

и множество  $K = \{ \alpha(e): e \in E_0 \}$  для конкретизации  $b_0$  готово к передаче и

ведет к последующей конкретизации  $b_1$ .

(3) Для каждого процесса  $p = (E_0, \leq_0, \alpha)$ , который не удовлетворяет условиям (1)

или (2), отношение

$$b_0 \xrightarrow{p} b_2$$

справедливо только тогда, когда для каждого непустого префикс-процесса

$p_1 = (E_1, \leq_1, \alpha_1)$  со свойством

$$p_1 \subseteq p \wedge p_1 \neq p$$

существует такая конкретизация  $b_1$ , что для процесса  $p_2 = p \setminus$

$E_0 \setminus E_1$

справедливо следующее высказывание:

$$b_0 \xrightarrow{p_1} b_1 \wedge b_1 \xrightarrow{p_2} b_2$$

Это определение принимает во внимание в пункте (2) то обстоятельство, что вентили в сети Петри рассматриваются только как последовательные единицы. Параллельные события всегда маркируются различными вентилями.

$$a \longrightarrow c$$

$$b \longrightarrow e$$

Для каждого конечного процесса  $p$  и каждой конкретизации  $b_0$  через

$$b_0 \xrightarrow{p} b_1$$

последующая конкретизация определяется однозначно. Она получается как простой итог встречающихся действий и, соответственно, вентиляй, по отношению к их входным и выходным местам.

**Лемма.** Если для процесса  $p = (E_0, \leq_0, \alpha)$  и конкретизации  $b_0$  и  $b_1$  справедливо высказывание

$$b_0 \xrightarrow{p} b_1,$$

то для каждого префикса  $p_1$  процесса  $p$  с множеством событий  $E_1$  справедливо следующее высказывание: для каждого множества событий  $E_2$ , которое является минимальным относительно  $\leq_0$  в  $E_0 \setminus E_1$ , справедливо, что множество  $\{ \alpha(e): e \in E_2 \}$  является готовым к пере-

даче.

**Доказательство.** Подпроцесс в  $p$ , который состоит только из событий из  $E_2$ , для конкретизации  $b_1$  является ходом работы.

**Лемма.** Если процесс  $p$  есть ход работы сети  $N$  с исходной конкретизацией  $b_0$ , то каждый префикс-процесс  $p_1 \subseteq p$  есть ход работы сети  $N$  с исходной конкретизацией  $b_0$ .

**Доказательство.** Согласно определению понятия хода работы для сети Петри процесс  $p$  есть ход работы, если любой префикс  $p$  есть ход работы.

Структура действий  $p = (E_0, \leq_0, \alpha)$  называется *совершенным ходом работы сети Петри  $N$  с конкретизацией  $b_0$* , если справедливо одно из следующих условий:

(1) Множество событий  $E_0$  бесконечно, и любой конечный префикс  $p$  есть ход работы сети  $N$  с исходной конкретизацией  $b_0$ .

(2) Множество событий  $E_0$  конечно, и  $p$  есть ход работы сети  $N$  для начальной конкретизации  $b_0$  с конечной конкретизацией  $b_1$ , и для  $b_1$  не существует непустого готового к передаче множества.

По этому определению сеть Петри описывает для каждой заданной исходной конкретизации множества процессов, а именно ходы работы и совершенные ходы работы сети.

Две сети с исходными конкретизациями называются *эквивалентными* (с точки зрения их работы), если они имеют одинаковые множества совершенных ходов работы.

**Лемма (линеаризация процессов).** Если  $p$  есть ход работы сети Петри  $N$  с исходной конкретизацией  $b$ , то каждая линеаризация  $p$  также есть ход работы  $N$  с исходной конкретизацией  $b$ .

**Доказательство.** Каждая линеаризация (по лемме линеаризации для готовых к передаче множеств) может через линеаризацию достичь определенных готовых к переходу множеств.

Обратное утверждение леммы не справедливо. Из множества последовательных ходов работы сети Петри нельзя сделать вывод о множестве не последовательных ходов работы.

Бесконечный ход работы  $p = (E_0, \leq_0, \alpha)$  сети Петри  $N$  с исходной конкретизацией  $b$  называется *несправедливым по отношению вентиля  $a$* , если одновременно выполняются следующие условия:

(1) Множество  $\{e \in E_0: \alpha(e) = a\}$  событий, помеченных действием  $a$ , конечно.

(2) Множество конечных префиксных процессов  $p_1 \subseteq p$ , при-

чем  $p_1 = (E_1, \leq_1, \alpha_1)$  с

$$\{e \in E_0: \alpha(e) = a\} = \{e \in E_1: \alpha_1(e) = a\}$$

и

$$b_0 \xrightarrow{p} b_1,$$

причем вентиль  $a$  в  $b_1$  готов к передаче, является бесконечным.

В несправедливом по отношению действия  $a$  бесконечном ходе работы сети Петри вентиль  $a$  встречается лишь конечное число раз, хотя он бесконечно часто готов к передаче. Ход работы называется справедливым, если он не является несправедливым относительно никаких вентилях.

Описанное понятие справедливости является лишь одним из вариантов среди многих понятий подобного рода.

### 6.2.2. Механизм агентов

Для моделирования поведения дискретных систем могут быть привлечены множества процессов. Такие множества допускают также описания с помощью термов языков программирования. Это можно продемонстрировать на примере простого языка.

Для представления систем используются языковые термы, называемые *агентами*. Рассмотрим следующий БНФ-синтаксис языка агентов:

```

<agent> ::= skip
<action>
<agent>; <agent>
<agent> or <agent>
<agent> || <agent>
<agent> id
<agent id>:: <agent>

```

В данном случае  $\langle \text{action} \rangle$  обозначает некоторый наперед заданный принятый язык для описания отдельных действий, а  $\langle \text{agent id} \rangle$  - наперед заданное множество идентификаторов для агентов. С помощью этого языка агентов, как и с помощью сетей Петри, могут описываться множества процессов. Эти множества снова означают ходы работ агентов. Агент

skip

может выполнить только пустой процесс.

Агент, который состоит только из задания действия  $a$ , в качестве ходов работ имеет множество процессов, которые обладают единственным событием, помеченным действием  $a$ . Агент

$t_1; t_2$

соответствует последовательной композиции процессов, описанных агентами  $t_1$  и  $t_2$ . Агент

$t_1$  or  $t_2$

в качестве множества процессов обладает объединением множеств процессов, описанных агентами  $t_1$  и  $t_2$ . Агент

$t_1 \parallel t_2$

описывает множество процессов, которое содержит параллельную композицию некоординируемых процессов, описываемых агентами  $t_1$  и  $t_2$ . Пусть  $t$  есть терм агента, в который свободно входит идентификатор агента  $x$ .

Агент соответствует агенту (вызову агента)  $x$ , который характеризуется рекурсивным объявлением  $x \equiv t$ . С помощью рекурсивно определенных агентов могут быть специфицированы бесконечные процессы.

При известных обстоятельствах необходимо использовать неименованные вспомогательные вентили, чтобы выразить поведение агента сетью Петри.

Для каждого термина  $t$  синтаксической единицы  $\langle \text{agent} \rangle$  и каждой структуры действий

$p = (E_0, \leq_0, \alpha)$

можно аналогично конкретизациям сетей Петри установить, представляет ли структура действий  $p$  возможный ход работы агента  $t$ . Это можно сделать формально с помощью определения отношения (аналогично отношениям переходов между состояниями сети Петри)

$t_1 \xrightarrow{p} t_2$ .

которое свидетельствует о том, что агент  $t_1$  может выполнить процесс  $p$  и после этого ведет себя как агент  $t_2$

Это отношение устанавливает какие ходы работы возможны для агентов и к каким агентам они ведут, определяется рекурсивно следующим образом:

(1) Каждый агент может выполнить пустой процесс. Для пустого процесса  $p_0$  и любого агента  $t$  определим соответственно:

$t \xrightarrow{p_0} t$ .

(2) Агент  $a$ , состоящий из единственного действия  $a$ , может выполнить процесс, который содержит единственное событие, помеченное действием  $a$ .

$a \xrightarrow{p} \text{skip}$ ,

(3) Пусть  $p, p_0, p_1$  и  $p_2$  - любые процессы. Агент  $t_1$  or  $t_2$  показы-

ведет неде-терминированно поведение агента  $t_1$  или агента  $t_3$ .

$$t_1 \xrightarrow{p} t_2 \Rightarrow t_1 \text{ or } t_2 \xrightarrow{p} t_3,$$

$$t_2 \xrightarrow{p} t_3 \Rightarrow t_1 \text{ or } t_2 \xrightarrow{p} t_3.$$

Агент  $t_1; t_2$  ведет себя как агент  $t_1$  и, если  $t_1$  этим поведением переводится в **skip**, в заключение ведет себя как агент  $t_2$ . Соответственно определяются следующие два правила для последовательной композиции агентов:

$$t_1 \xrightarrow{p} t_1 \Rightarrow t_1; t_2 \xrightarrow{p} t_3; t_2,$$

$$t_1 \xrightarrow{p_1} \text{skip} \wedge t_2 \xrightarrow{p_2} t_3 \wedge \text{isseq}(p, p_1, p_2) \Rightarrow t_1; t_2 \xrightarrow{p} t_3$$

Поведение агента  $t_1 \parallel t_2$  складывается параллельно из поведения  $t_1$  и  $t_2$ . Если поведение обоих агентов приводит к **skip**, то и поведение  $t_1 \parallel t_2$  тоже приводит к **skip**. Агент  $t_1 \parallel t_2$  завершается, если завершается как  $t_1$ , так и  $t_2$ . Соответственно определяем:

$$t_1 \xrightarrow{p_1} r_1 \wedge t_2 \xrightarrow{p} r_2 \wedge \text{ispar}(p, p_1, p_2, \emptyset) \Rightarrow t_1 \parallel t_2 \xrightarrow{p} r_1 \parallel r_2,$$

$$t_1 \xrightarrow{p_1} \text{skip} \wedge t_2 \xrightarrow{p} \text{skip} \wedge \text{ispar}(p, p_1, p_2, \emptyset) \Rightarrow t_1 \parallel t_2 \xrightarrow{p} \text{skip}.$$

(4) Для любого процесса  $p$  определим ходы работ для рекурсивных агентов с помощью следующего правила:

$$\{ [ x::t/x ] \xrightarrow{p} t_0 \Rightarrow x::t \xrightarrow{p} t_0$$

Теперь через определенное выше отношение каждому агенту можно предписать множество процессов. Процесс  $p$  называется (не-совершенным) ходом работы агента  $t_0$ , если существует агент  $t_1$

$$t_0 \xrightarrow{p} t_1.$$

Множество ходов работ является замкнутым относительно образования префиксов.

Лемма (закон разложения). Если для агентов  $t_0, t_2$  и конечного процесса  $p$  справедливо высказывание

$$t_0 \xrightarrow{p} t_2,$$

то для каждого префикса  $p_1 \subseteq p$  существует агент  $t_1$  такой, что

$$t_0 \xrightarrow{p_1} t_1, t_1 \xrightarrow{p_1} t_2,$$

Доказательство индукцией по числу событий в  $p_1$ .

Закон разложения означает, что осуществляемые с помощью процессов переходы также могут быть разложены в следующие друг за другом переходы, причем эти последние осуществляются подпро-

цессами исходного процесса. Справедливо также и обратное утверждение.

Предложение (связывание процессов). Если для агентов  $t_1, t_2, t_3$  и процессов  $p_0, p_1$  справедливо высказывание

$$t_1 \xrightarrow{p_0} t_2 \wedge t_2 \xrightarrow{p_1} t_3,$$

то существует процесс  $p_2$  такой, что

$$t_1 \xrightarrow{p_2} t_3,$$

Доказательство индукцией по структурам агентов.

Агент  $t_0$  называется *терминальным*, если для любого агента  $t_1$  высказывание

$$t_0 \xrightarrow{p} t_1$$

справедливо только с пустым процессом  $p$ . Таким образом, терминальный агент может осуществлять переходы только с пустым процессом, т. е. не выполняя каких-либо действий.

Процесс  $p$  называется *совершенным ходом работы* агента  $t_0$ , если для терминального агента  $t_1$  справедливо:

$$t_0 \xrightarrow{p} t_1$$

или же процесс  $p$  является бесконечным и каждый его префикс  $p_1 \subseteq p$  является ходом работы.

Два агента называются *эквивалентными*, если они обладают одинаковыми множествами совершенных ходов работы.

К анализу агентов можно поставить такие же вопросы, какие ставятся и для сетей Петри. Концепция сети Петри с конкретизациями заменяется концепцией агентов. Тем самым агенту соответствует при сетях пара  $(N, b)$ , причем  $N$  представляет сеть, а  $b$  - ее конкретизацию. В случае сетей только рассматривается отношение переходов на конкретизациях, так как сеть оставляет переходы неизменными.

Для агента  $t_0$  агент  $t_1$  называется *достижимым*, если существует процесс  $p$  такой, что

$$t_0 \xrightarrow{p} t_1$$

Для агента, который моделирует распределенную систему или определенный ее аспект, можно сформулировать ряд вопросов:

1. Какими свойствами обладает множество достижимых агентов? Является ли это множество конечным?
2. Исключается ли для определенных действий, что будет порожден процесс, в котором оба эти действия имеют место одновременно (взаимное исключение)?
3. Возможно ли быть с помощью процесса достигнут агент



(отличный от *skip*), для которого возможны только тривиальные переходы (тупики)?

4. Может ли быть достигнут агент (отличный от *skip*), исходя из которого больше не достигим никакой агент, в котором имеют место определенные действия (локальный тупик, отсутствующая живучесть)?

5. Существуют ли бесконечные процессы, в которых одно из действий никогда не встречается (справедливость)?.

### Синхронизация и координация агентов

При параллельной композиции агентов не накладываются никакие ограничения на отношения порядка между событиями агентов. Однако часто возникает необходимость показать, что между множествами событий процессов параллельно протекаемых агентов могут иметь место определенные отношения. В таком случае следует рассматривать *координацию* или *синхронизацию* параллельных процессов. Для этой цели наряду с уже представленной полностью независимой параллельной композицией агентов вводится оператор композиции, который разрешает координацию параллельно выполняющихся агентов.

Координация нужна для того, чтобы при параллельно протекающих процессах можно было исключить конфликты.

Пусть  $S$  - подмножество множества действий  $A$ . Процесс  $p_0 = (E_0, S_0, \alpha_0)$  может быть сложен из процессов  $p_1$  и  $p_2$  тем, что события, помеченные действиями из  $S$ , установлены в качестве общих событий. Тогда справедливо:

$$\text{ispar}(p_0, p_1, p_2, S) .$$

Рассмотрим частный случай для установления общих событий. Множество действий  $A$  разлагается на два дизъюнктивных подмножества  $U$  и  $S$ :

действия множества  $U$  являются асинхронно выполняемыми действиями,

действия множества  $S$  являются только синхронно выполняемыми действиями..

Введение параллельных композиций агентов, которые относительно событий, помеченных определенными действиями, должны протекать синхронно, порождает новые вопросы и проблемы. Что произойдет, если оба параллельно выполняющиеся агента настаивают на синхронном выполнении *различных* (не согласующихся друг с другом) действий? В этом случае не может иметь место событие (не выполняется никакое действие), поскольку для каждой пары син-

хронных действий требовалась бы готовность обоих агентов. Возникает ситуация *тупика*.

Наряду с несинхронизируемой параллельной композицией агентов важно рассмотреть и синхронизируемую параллельную композицию агентов. Это может происходить посредством синхронной параллельной композиции  $\parallel_s$ , где  $S$  обозначает множество действий, которые при параллельно протекающих процессах должны выполняться синхронно.

Точное значение синхронной параллельной композиции по отношению к ходу течения процесса задается следующими правилами:

$$\begin{aligned}
 t_1 \xrightarrow{p_1} r_1 \wedge t_2 \xrightarrow{p} r_2 \wedge \text{ispar}(p, p_1, p_2, \emptyset) &\Rightarrow t_1 \parallel_s t_2 \\
 &\xrightarrow{p} r_1 \parallel r_2, \\
 t_1 \xrightarrow{p_1} \text{skip} \wedge t_2 \xrightarrow{p} \text{skip} \wedge \text{ispar}(p, p_1, p_2, \emptyset) &\Rightarrow t_1 \parallel_s t_2 \\
 &\xrightarrow{p} \text{skip}.
 \end{aligned}$$

Эти правила являются обобщением рассмотренных выше правил для несинхронизируемой параллельной композиции. Здесь имеет место то, что какой-либо процесс есть ход работы агента  $t_1 \parallel_s t_2$ , если он возникает через параллельную композицию ходов работы  $t_1$  и  $t_2$ , причем в параллельной композиции этих процессов события, помеченные действиями из  $S$ , являются общими событиями. По этому определению существуют агенты, которые отличны от *skip*, но тем не менее являются терминальными и тем самым не могут выполнять какого-либо процесса кроме пустого (такие агенты находятся в тупике).

**Пример (тупик).** Агент

$$a ; b \parallel_{\{a, b\}} b ; a$$

находится в тупике, поскольку он не может выполнять ни действия  $a$ , ни действия  $b$ .

Наряду с синхронизацией процессов относительно определенных действий над общими событиями имеет место координация параллельно выполняющихся агентов о «взаимных исключениях». Определенные последовательности действий одного агента должны выполняться таким образом, чтобы параллельно не происходили определенные действия другого агента. Иначе возникнут конфликты между действиями агентов. С помощью синхронной параллельной композиции

$$t_1 \parallel_s t_2,$$

порождается процесс, который (при отсутствии возникновения тупиков) содержит определенные линеаризации ходов работы  $t_1$  и  $t_2$ ,

как частичных процессов. Возникают дальнейшие причинные отношения между событиями в обоих частичных процессах - это следует ввести для обеспечения взаимных исключений.

**Пример (взаимное исключение через пару синхронизируемых действий).** Пусть даны следующие три агента:

$$t_1 = x_1 :: a_1 ; p ; b_1 ; v ; x_1$$

$$t_2 = x_2 :: a_2 ; p ; b_2 ; v ; x_2$$

$$k = y :: p ; v ; y.$$

Предположим, что действия  $b_1$  и  $b_2$  могут конфликтовать между собой и следовательно не должны выполняться параллельно. Для агента

$$t_1 \parallel t_2$$

возможен ход работы, заданный с помощью следующего графа действий:

$$a_1 \rightarrow p \rightarrow b_1 \rightarrow v \rightarrow a_1 \rightarrow p \rightarrow b_1 \rightarrow v \rightarrow \dots$$

$$a_2 \rightarrow p \rightarrow b_2 \rightarrow v \rightarrow a_2 \rightarrow p \rightarrow b_2 \rightarrow v \rightarrow \dots$$

Этот процесс содержит пару параллельных событий, которые помечены конфликтующими действиями  $b_1$  и  $b_2$ . В агенте

$$(t_1 \parallel t_2) \parallel \{v ; p\}k$$

такой ход работы исключается. Каждое выполнение действий  $p$  и  $v$  агента  $k$  совпадает с одними из действий  $p$  и  $v$  или агента  $t_1$ , или агента  $t_2$ , но что  $p$  и  $v$  - действия в процессах  $t_1$  и  $t_2$  не помечают каких-либо общих событий.

Методика обеспечения взаимных исключений, описанная в примере для агентов, может быть использована в более общих случаях.

Пусть  $a$  - действие, находящееся в конфликте с самим собой, которое встречается сколь угодно часто в агентах  $t_1$  и  $t_2$ . В агентах  $t_1$  и  $t_2$  заменим действие  $a$  на терм агентов

$$P ; a ; v,$$

причем пусть  $p$  и  $v$  - до сих пор не встречавшиеся действия в  $t_1$  и  $t_2$ . Агент  $c :: p ; v ; c$  в каждом ходе работы агента по мере надобности вынуждает один из обоих агентов  $t_i'$  или  $t_i$  ждать с выполнением  $p ; a ; v$ -до тех пор, пока агент  $c :: p ; v ; c$  снова будет готов для события, помеченного действием  $p$ . События  $c$  с действием  $a$  в  $t_1'$  или  $t_2'$  в ходах работы, описанных с помощью агента

$$(t_1' \parallel t_2') \parallel_{(p, v)} c :: p ; v ; c,$$

полностью линейно упорядочены. При таком способе действий для обеспечения взаимных исключений возникают агенты, которые обнаруживают дополнительные возможные тупики - в случае, когда

в агенте  $t_1$  или  $t_2$  в параллельных композициях действие  $a$  должно выполняться синхронно.

**Лемма (взаимное исключение).** Для любого агента  $t$ , в котором действие  $a$  всегда встречается в виде  $r$ ;  $a$ ;  $v$  и иначе действия  $r$  и  $v$  не встречаются, справедливо, что каждый из процессов, порождаемых агентом не содержит никакой пары параллельных событий, помеченных действием  $a$ .

**Доказательство.** Пусть  $r'$  - процесс, порождаемый агентом  $t'$ . События, которые помечены действиями  $r$  или  $v$ , образуют последовательный подпроцесс, в котором  $r$  и  $v$  чередуются. Каждому событию  $e$ , помеченному через  $a$ , может быть однозначным образом предписано натуральное число событий, которые являются причиной для  $e$  и которые помечены через  $v$ , соответственно через  $r$ . Каждому событию  $e$ , помеченное через  $a$ , можно однозначно предписать помеченное через  $r$  событие  $er_r(e)$ , которое является причиной для  $e$ , а также помеченное через  $v$  событие  $ev_v(e)$ , для которого  $e$  является причиной.

### 6.2.3. Предикаты над процессами

Наряду с сетями Петри и терминами агентов множества процессов могут быть охарактеризованы через задание свойств с помощью выражений логики предикатов. Мы будем обсуждать лишь очень простые предикаты.

Пусть задан процесс  $p = (E_0, \leq_0, \alpha)$ , с помощью термина  $\#(a, p)$

обозначим число событий в  $p$ , помеченных действием  $a$ . Точно  $\#(a, p)$  определяется следующей формулой:

$$\#(a, p) = |\{e \in E_0: \alpha(e) = a\}|.$$

Для бесконечных процессов значение  $\#(a, p)$  может быть бесконечным. С помощью этой формулы можно образовывать предикаты и тем самым описывать множества процессов.

**Пример (множества процессов через предикаты).** Рассмотрим следующее множество  $A$  действий:

$$A = \{a, b, c, d\}.$$

Множество конечных процессов  $p = (E_0, \leq_0, \alpha)$ , можно охарактеризовать через предикат  $Q(p)$ . Пусть  $Q(p)$  задан с помощью следующих формул:

$$(1) 0 \leq \#(a, p) - \#(c, p) \leq 1 \wedge$$

$$(2) 0 \leq \#(b, p) - \#(d, p) \leq 1 \wedge$$

$$(3) \forall e, e' \in E_0: \alpha(e) - c \wedge \alpha(e') = d \Rightarrow (e \leq_0 e' \vee e' \leq_0 e).$$

Процесс  $p$  выполняет предикат  $Q$ , если:

- (1) число событий, помеченных через  $a$ , либо на единицу больше числа событий, помеченных через  $c$ , либо равно ему;
- (2) Число событий, помеченных через  $b$  либо на единицу больше числа событий, помеченных через  $d$ , либо равно ему;
- (3) События, помеченные через  $c$  и  $d$  не протекают параллельно.

### 6.3. Вычислимость в геоинформатике

Не все постановки задач информационной обработки могут быть решены с помощью алгоритмов. Точнее говоря, существуют задачи, которые могут быть математически точно описаны («формализуемы»), но для решения которых не существует алгоритмов. Здесь рассматривается задача вычисления некоторой определенной функции. Эта задача решается по алгоритму, который вычисляет функцию, описанную в постановке задачи. Если существует алгоритм для вычисления некоторой функции, то она вычислима.

Итак, функция  $f$  (которая определена постановкой задачи) называется *вычислимой*, если существует алгоритм, который для любого аргумента  $x$  (в качестве входа для алгоритма) вычисляет значение  $f(x)$  (как выход алгоритма). Тем самым понятие вычислимости сводится к понятию алгоритма. В частности, высказывание о вычислимости функции зависит от выбора понятия алгоритма.

Существует несколько различных подходов к формализации понятия алгоритма. Некоторые выбранные из этих подходов обсуждаются в настоящей работе и показано, что предложенные до сих пор понятия алгоритма приводят к одинаковому понятию вычислимости.

Здесь вычисляются  $n$ -местные функций  
 $f: N^n \rightarrow N$

над натуральными числами. Вводимые и исследуемые в дальнейшем понятия вычислимости распространяются на любые области определений и значений функций.

Можем производить вычисления только на представлениях элементов из  $N$ , а не на самих этих элементах, для представления чисел будут использоваться множества  $T$  знаков и инъективное отображение

$гер: N \rightarrow T^*$ ,

которое устанавливает представление чисел в виде слов над  $T$ . Предполагается, что отображение  $гер$  обратимо на свой прообраз, поскольку только в этом случае получается однозначное представление чисел. В соответствии с этим существует отображение

$\text{abs}: \{t \in T^*: \exists n \in N : \text{rep}(n) = t\} \rightarrow N$ ,  
 так что для отображений  $\text{rep}$  и  $\text{abs}$  справедливы равенства  
 $\text{abs} \circ \text{rep} = \text{id}$ ,  
 $\text{rep} \circ \text{abs} = \text{id}$ .

Эта методика перевода конкретных представлений в множество абстрактных понятий с помощью функции абстракции  $\text{abs}$  типична для образа действий специалиста по информатике геоинформатике.

С помощью этого перевода вместо абстрактного отображения между натуральными числами

$f: N'' \rightarrow N$

рассматривается конкретное отображение между последовательностями знаков

$f: (T^*) \gg \rightarrow T^*$

со следующим свойством (для всех  $x_1, \dots, x_n \in N$ ):

$f(x_1, \dots, x_n) = \text{abs}(f(\text{rep}(x_1), \dots, \text{rep}(x_n)))$ .

Это равенство говорит, что, вместо того чтобы вычислять значение функции  $f$  для аргументов  $x_1, \dots, x_n$ , можно определить представления  $\text{rep}(x_1), \dots, \text{rep}(x_n)$  аргументов и для них вычислить значение  $f$ . С помощью функции  $\text{abs}$  можно из результата применения  $f$  к  $\text{rep}(x_1), \dots, \text{rep}(x_n)$  получить значение функции для аргументов  $(x_1, \dots, (x_n)$ . Из приведенного выше равенства получается *коммутирующая диаграмма*, показанная на рис. 6.3.

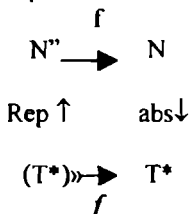


Рис. 6.3. Коммутирующая диаграмма

Эту схему отношений (между абстрактным и конкретным представлениями) можно найти во многих областях пошагового уточнения представления данных в программировании и относящихся к этому функций. В таких случаях говорится об *аспекте абстракции*.

Употребительными формами представления чисел являются представление соответствующим количеством штрихов, двоичное или десятичное представления и разложением на простые числа. Существует зависимость между сложностью вычислений и выбором

конкретного представления. Для определения понятия вычислимости дополнительно требуется, чтобы функции  $\text{ger}$  и  $\text{abs}$  сами были вычислимы в интуитивном смысле.

Конкретные алгоритмы работают всегда с представлениями натуральных чисел. Поэтому необходимо, чтобы конкретные представления были также реалистически приемлемы для обращения. При данном условии предполагается, что  $T$  есть конечное множество.

#### 6.4. Фракталы как метод описания пространственных данных

Понятия фрактал и фрактальная геометрия, появившиеся в конце 70-х, с середины 80-х прочно вошли в обиход математиков и программистов. Термин фрактал образован от латинского *fractus* (состоящий из фрагментов). Он был предложен Бенуа Мандельбротом в 1975 году для обозначения нерегулярных, но самоподобных структур, которыми он занимался. Рождение фрактальной геометрии принято связывать с выходом в 1977 году книги Мандельброта *'The Fractal Geometry of Nature'*. В его работах использованы научные результаты других ученых (Пуанкаре, Фату, Жюлиа, Кантор, Хаусдорф), работавших в период 1875-1925 годов в той же области. Но только в наше время удалось объединить их работы в единую систему.

Основное свойство фракталов самоподобие или рекурсивность. Оно является одновременно достоинством и ограничением применения этого подхода. Самоподобие позволяет в небольшой части фрактала содержать информацию о всем фрактале.

Применение самоподобия позволяет осуществлять связь большого с малым или малого с большим, что широко используется в геоинформатике в виде процедур генерализации или детализации.

В процессе генерализации или обобщения возникает задача построения более мелкого графического объекта на основе более крупного при обобщении мелких деталей и сохранении характерных крупных деталей. Другими словами, ставится задача перехода к новому графическому образу с сохранением связи с прежним образом. Фрактальный подход позволяет решать эту задачу на основе принципа самоподобия.

Для пояснения фрактальной размерности вводят понятие топологической размерности. Под топологической размерностью  $D_1$  множества в линейном пространстве понимают число линейно независимых координат в пространстве. Например, окружность и линия

имеют топологическую размерность 1; круг и квадрат — 2; шар и куб — 3.

**Фрактальная размерность** множества  $D$  — размерность того пространства, которое полностью заполняется множеством. Для связи фрактальной и топологической размерностей используют показатель Херста  $H$ , вычисляемый по формуле:  $H = D - Dt$ .

**Фракталом** называют множество, фрактальная размерность которого не совпадает с топологической.

Велика роль фракталов в машинной графике. Они применяются, например, когда требуется с помощью нескольких коэффициентов задать линии и поверхности очень сложной формы. С точки зрения машинной графики фрактальная геометрия незаменима при генерации искусственных облаков, гор, поверхности моря. Фактически найден способ легкого представления сложных неевклидовых объектов, образы которых весьма *похожи* на природные. Это свойство является актуальным именно в геоинформатике.

Фракталы в зависимости от области их построения делятся на разные классы, например: геометрические, алгебраические, стохастические.

Геометрические фракталы представляют собой пространственные образования, создающие взаимоподобные геометрические фигуры.

Алгебраические образуют самую крупную группу фракталов. Получают их с помощью нелинейных процессов в  $n$ -мерных пространствах. Наиболее изучены двухмерные процессы. Интерпретируя нелинейный итерационный процесс как дискретную динамическую систему, можно пользоваться терминологией теории этих систем: *фазовый портрет, установившийся процесс, аттрактор* и т.д.

Нелинейные динамические системы обладают несколькими устойчивыми состояниями. Как правило, состояние, в котором оказалась динамическая система после некоторого числа итераций, зависит от ее начального состояния. Поэтому каждое устойчивое состояние (или как говорят - аттрактор) обладает некоторой областью начальных состояний, из которых система обязательно попадет в рассматриваемые конечные состояния.

Таким образом, фазовое пространство сложной системы разбивается на области устойчивости или *области притяжения* аттракторов. Если фазовым является двухмерное пространство, то окрашивая области притяжения различными цветами, можно получить *цветовой фазовый портрет* этой системы (итерационного процесса).



Изменяя алгоритм выбора цвета, можно получить сложные фрактальные картины с причудливыми многоцветными узорами. Этот подход создает возможность с помощью примитивных алгоритмов порождать очень сложные нетривиальные структуры.

В качестве примера часто используют множество Мандельброта. Алгоритм его построения основан на простом итеративном выражении:

$$Z[i+1] = Z[i] * Z[i] + C,$$

где  $Z_i$  и  $C$  - комплексные переменные. Итерации выполняются для каждой стартовой точки  $C$  прямоугольной или квадратной области - подмножестве комплексной плоскости. Итерационный процесс продолжается до тех пор, пока  $Z[i]$  не выйдет за пределы окружности радиуса 2, центр которой лежит в точке  $(0,0)$ , (это означает, что аттрактор динамической системы находится в бесконечности), или после достаточно большого числа итераций (например 200-500)  $Z[i]$  сойдется к какой-нибудь точке окружности. В зависимости от количества итераций, в течении которых  $Z[i]$  оставалась внутри окружности, можно установить цвет точки  $C$  (если  $Z[i]$  остается внутри окружности в течение достаточно большого количества итераций, итерационный процесс прекращается и эта точка раstra окрашивается в черный цвет).

Вышеописанный алгоритм дает приближение к так называемому множеству Мандельброта. Множеству Мандельброта принадлежат точки, которые в течение *бесконечного* числа итераций не уходят в бесконечность (точки имеющие черный цвет). Точки принадлежащие границе множества (именно там возникает сложные структуры) уходят в бесконечность за конечное число итераций, а точки лежащие за пределами множества, уходят в бесконечность через несколько итераций (белый фон).

Еще одним известным классом фракталов являются стохастические фракталы. Этот вид основан на получении самоподобных совокупностей, которые строятся, при условии, что в итерационном процессе случайным образом изменяются какие-либо его параметры. При этом получаются объекты очень похожие на природные - несимметричные деревья, изрезанные береговые линии и т.д. Двумерные стохастические фракталы используются при моделировании рельефа местности и поверхности моря [2].

Существуют и другие классификации фракталов, например деление фракталов на детерминированные (алгебраические и геометрические) и недетерминированные (стохастические).

В качестве примера построения фрактальных совокупностей вернемся к геометрическим фракталам.

Рассмотрим классический пример фрактала - триадную кривую Коха и шаги его построения.

Шаг 1. Построение кривой начинается с единичного отрезка, который называется инициатором (генератором) и является предфракталом 0-го порядка



Рис.6.4. Предфрактал 0-го порядка

Шаг 2. Инициатор заменяется на *образующий элемент* - в нашем случае кривую из четырех звеньев, каждое из которых имеет длину  $1/3$ . Так образуется предфрактал 1-го порядка



Рис.6.5. Предфрактал 1-го порядка

Длина полученного фрактала равна  $4/3$  от исходной

Шаг 3. Для построения предфрактала следующего порядка каждое звено заменяется на уменьшенный образующий элемент. В результате получаем кривую, состоящую из  $4 \times 4 = 16$  звеньев (рис.6.6.), каждое из которых имеет длину  $(1/3)/3 = 1/9$ , общая длина равна  $16/9$ .

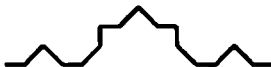


Рис.6.5. Предфрактал 2-го порядка

Длина предфрактала  $n$ -го порядка равна  $(4/3)^n$  в степени  $n$ . Очевидно, что предел длины кривой при  $n$ , стремящемся к бесконечности, равен бесконечности.

Продолжая деление отрезков по заданному принципу можно строить новые фигуры других масштабов, которые, с одной стороны, будут отличаться от исходных, с другой, содержать очертания и

главное метод перехода от фигуры одного масштаба к фигуре другого масштаба.

Если построение кривой начинать не с отрезка, а с треугольника, и применить вышеперечисленные построения к каждой его стороне, то получим «снежинку» Кох. Эта фигура интересна тем, что ее периметр — линия бесконечной длины — ограничивает конечную площадь.

Фракталы используют также для описания трехмерных объектов. Главным принципом является задание рекурсивного преобразования и исходной фигуры - инициатора.

Рассмотренный эвристический подход имеет аналитическое обоснование, известное под названием метод "Систем Итерированных Функций" (Iterated Functions System - IFS), который появился в середине 80-х годов как средство получения фрактальных структур.

IFS представляет собой систему функций из некоторого фиксированного класса функций, отображающих одно многомерное множество на другое. Наиболее простая IFS состоит из аффинных преобразований плоскости:

$$\begin{aligned} X' &= A * X + B * Y + C \\ Y' &= D * X + E * Y + F \end{aligned}$$

В 1988 году известные американские специалисты в теории динамических систем и эргодической теории Барнсли и Слоан предложили некоторые идеи, основанные на соображениях теории динамических систем, для сжатия и хранения графической информации. Они назвали свой метод методом фрактального сжатия информации. Происхождение названия связано с тем, что геометрические образы, возникающие при использовании этого метода, обычно имеют фрактальную природу в смысле Мандельброта.

На основании этих идей Барнсли и Слоан создали алгоритм, который, по их утверждению, позволит сжимать информацию в 500-1000 раз.. Вкратце метод можно описать следующим образом. Изображение кодируется несколькими простыми преобразованиями (в нашем случае аффинными), т.е. коэффициентами этих преобразований (в нашем случае A, B, C, D, E, F).

Например, закодировав какое-либо изображение двумя аффинными преобразованиями, мы однозначно определяем его с помощью 12-ти коэффициентов. Если после этого задаться какой-либо начальной точкой (например  $X=0$   $Y=0$ ) и запустить итерационный процесс, то в результате первой итерации получим две точки, после второй - четыре, после третьей - восемь и т.д. Через несколько десятков ите-

раций совокупность полученных точек будет описывать закодированное изображение. Трудность этого подхода заключается в том, что очень трудно найти коэффициенты IFS, которая кодировала бы произвольное изображение.

Для построения IFS применяют кроме аффинных и другие классы простых геометрических преобразований, которые задаются небольшим числом параметров. Например, проективные:

$$X' = (A1 * X + B1 * Y + C1) / (D1 * X + E1 * Y + F1)$$

$$Y' = (A2 * X + B2 * Y + C2) / (D2 * X + E2 * Y + F2)$$

или квадратичные:

$$X' = A1 * X * X + B1 * X * Y + C1 * Y * Y + D1 * X + E1 * Y + F1$$

$$Y' = A2 * X * X + B2 * X * Y + C2 * Y * Y + D2 * X + E2 * Y + F2$$

преобразования на плоскости.

Таким образом фрактальная геометрия дает возможность осуществлять кодирование изображения, его сжатие, осуществлять генерализацию или детализацию пространственных объектов. По этой причине его применение в геоинформатике перспективно.

## 6.5. Обработка графической информации

### 6.5.1. Графические изображения

Преобразование и кодирование графической информации в компьютерную (цифровую) форму производится с помощью сканера.

Возрастающее значение имеет видеoinформация (телевизионные передачи и видеофильмы), преобразование которых в компьютерную форму производится с помощью специальных устройств: видеотюннеров и видеоадаптеров.

С помощью графических редакторов можно создать графическое изображение и непосредственно на компьютере (нарисовать картинку, создать чертеж, диаграмму, построить график и т.д.).

Графическая информация хранится и обрабатывается в компьютере в цифровой форме, т.е. в виде последовательностей нулей и единиц.

Все компьютерные изображения разделяют на два типа: растровые и векторные. Их различают по способу кодирования и представления.

Растровое изображение это упорядоченное изображение, является аналогом пространственной матрицы и хранится с помощью пикселей, которые образуют строки и столбцы. Каждый пиксель имеет определенное положение и цвет. Растровое изображение фор-

мируется из объектов одного класса - пикселей.

Качество растрового изображения зависит от числа пикселей на элемент реального изображения и количества цветов, которые могут принимать пиксели. Хранение каждого пикселя требует определенного количества бит (глубина цвета), которое зависит от количества цветов в изображении.

Растровые графические изображения обычно имеют большой размер и большую глубину цвета (24 или 36 бит на точку). В результате файлы, хранящие растровые изображения, имеют большой информационный объем.

Растровые изображения очень чувствительны к масштабированию (увеличению или уменьшению). При уменьшении растрового изображения несколько соседних точек преобразуются в одну, поэтому теряется разборчивость мелких деталей изображения. При увеличении — увеличивается размер каждой точки и появляется ступенчатый эффект, который можно увидеть невооруженным глазом.

Векторные изображения формируются из графических объектов разных классов (точка, линия, окружность, прямоугольник и т.д.), которые хранятся в памяти компьютера в виде графических примитивов и описывающих их математических формул.

Например, графический примитив точка задается своими координатами  $(X, Y)$ ; линия — координатами начала  $(X_1, Y_1)$  и конца  $(X_2, Y_2)$ ; окружность — координатами центра  $(X, Y)$  и радиусом  $(R)$ ; круг — координатами центра  $(X, Y)$  радиусом  $(R)$ , типом заполнения, цилиндр — координатами центра  $(X, Y)$ , радиусом  $(R)$ , высотой  $(H)$  и т.д. Для каждого примитива задается тип линии, толщина, цвет.

Векторные графические изображения являются удобным средством получения и хранения высокоточных графических объектов. Векторная графика является основой компьютерного черчения, автоматизированного проектирования, автоматизированного картографирования, трехмерного моделирования реальности.

Файлы, хранящие векторные изображения, имеют сравнительно меньший объем по сравнению с растровыми аналогами. Достоинством этого типа изображений является возможность увеличения или уменьшения без потери качества. Это возможно, т.к. масштабирование изображений производится с помощью простых математических операций (умножения параметров графических примитивов на коэффициент масштабирования).

Выше рассматривалось кодирование графических объектов и

было отмечено, что тип кодировки определяет формат графического файла, содержащего изображение. Форматы графических файлов определяют способ хранения информации в файле (растровый или векторный), а также форму хранения информации (используемый алгоритм сжатия).

Сжатие применяется для растровых графических файлов, т.к. они имеют обычно достаточно большой объем.

Сжатие графических файлов отличаются от их архивации с помощью программ-архиваторов (RAR, ZIP, ARJ и т.д.) тем, что алгоритм сжатия включается в формат графического файла. Существуют различные алгоритмы сжатия, которые целесообразно применять для различных типов изображения.

Сжатие рисунков, содержащих большие области однотонной закрашки, эффективно при использовании алгоритма сжатия, который заменяет последовательность повторяющихся величин (пикселей одинакового цвета) на две величины (пиксель и количество его повторений). Такой алгоритм сжатия используется в графических файлах форматов BMP и PCX.

Для рисунков типа диаграммы целесообразно применение другого метода сжатия, который использует поиск повторяющихся в рисунке «узоров». Этот метод отдаленно напоминает использование метода "самоподобия", служащий основой при построении фракталов. Такой алгоритм используется в графических файлах форматов TIF и GIF и позволяет сжать файл в несколько раз.

Для сжатия отсканированных фотографий и иллюстраций используется алгоритм сжатия JPEG. Этот алгоритм использует тот факт, что человеческий глаз очень чувствителен к изменению яркости отдельных точек изображения, но гораздо хуже замечает изменение цвета. Действительно, при глубине цвета 24 бита компьютер обеспечивает воспроизведение более 16 млн. различных цветов, тогда как человек вряд ли способен различить и тем более назвать более сотни цветов и оттенков.

Применение метода JPEG позволяет сжимать файлы в десятки раз, однако может приводить к необратимой потере информации (файлы не могут быть восстановлены в первоначальном виде). Следует отметить, что этот формат нецелесообразно применять при сканировании карт и высокоточных чертежей, так как он приводит к потере информации.

**Bit Map Image (BMP)** — универсальный формат растровых графических файлов, используемый во многих графических редак-

торах. Этот формат может содержать изображения с глубиной цвета 10, 16, 24 и 36 бит и рекомендуется для хранения и обмена данными с другими приложениями.

**Tagged Image File Format (TIFF)** — формат растровых графических файлов, поддерживаемый всеми основными графическими редакторами и компьютерными платформами. Этот формат включает в себя алгоритм сжатия без потери информации. Используется для обмена документами между различными программами. Рекомендуется для использования при работе с издательскими системами.

**Graphics Interchange Format (GIF)** — формат растровых графических файлов. Данный формат поддерживается приложениями для различных операционных систем. Включает алгоритм сжатия без потери информации, позволяющий уменьшить объем файла в несколько раз. Используется формат GIF для хранения изображений, создаваемых программным путем (диаграмм, графиков и т.д.), рисунков (типа аппликации) с ограниченным количеством цветов (до 256) и для размещения графических изображений на Web-страницах в Интернете.

**Portable Network Graphic (PNG)** — формат растровых графических файлов аналогичный формату GIF. Рекомендуется применять для размещения графических изображений на Web-страницах в Интернете.

**Joint Photographic Expert Group (JPG)** — формат растровых графических файлов, который реализует эффективный алгоритм сжатия (метод JPEG) для отсканированных фотографий и иллюстраций. Алгоритм сжатия позволяет уменьшить объем файла в десятки раз, однако приводит к необратимой потере части информации. Поддерживается данный формат приложениями для различных операционных систем, используется для размещения графических изображений на Web-страницах в Интернете.

**Windows MetaFile (WMF)** — упрощенный формат векторных графических файлов для Windows-приложений. Используется для хранения коллекции графических изображений Microsoft Clip Gallery.

**Encapsulated PostScript (EPS)** — формат векторных графических файлов, поддерживаемый программами для различных операционных систем. Обычно рекомендуется для печати и создания иллюстраций в настольных издательских системах.

**CorelDRaw files (CDR)** — оригинальный формат векторных графических файлов, используемый в системе обработки векторной

графики CorelDraw.

### 6.5.2. Графические редакторы

Для обработки изображений в геоинформационных технологиях используют специальные программы — *графические редакторы*. Графические редакторы разделяют на две категории: растровые и векторные по типу обрабатываемых изображений. Геоинформационных систем также делятся на два типа (растровые и векторные) и имеют соответствующие редакторы.

Растровые графические редакторы применяют для обработки фотографий и рисунков, поскольку они обеспечивают высокую точность передачи градаций цветов и полутонов.

Среди растровых графических редакторов есть простые, например, стандартное приложение Paint, и мощные профессиональные графические системы, например, Adobe Photoshop и CorelPhoto-Paint.

Векторные графические редакторы имеются во всех векторных ГИС и программах автоматизированного проектирования типа "Автокад".

Графические редакторы предоставляют возможность выбора инструментов для создания и редактирования графических изображений, объединяя их в *панели инструментов*. После выбора объекта на панели инструментов его можно нарисовать в любом месте окна редактора.

К примеру для рисования линии нужно выполнить следующие действия:

выбрать на панели инструментов инструмент *Линия*;

переместить курсор на определенное место окна редактора и щелчком мыши зафиксировать точку, из которой должна начинаться линия;

перетащить линию в нужном направлении и, осуществив повторный щелчок, зафиксировать второй край линии.

Следует подчеркнуть, что набор инструментов в растровом и в векторном графическом редакторах в основном совпадает, однако принципы работы с ними несколько различаются.

Так, в растровом графическом редакторе объект *Линия* после окончания рисования перестает существовать как самостоятельный элемент и становится лишь группой пикселей на рисунке. В векторном редакторе этот же объект продолжает сохранять свою индивидуальность и его можно масштабировать, перемещать по рисунку и т.д.



В векторном редакторе существует процедура группировки или построения блока. Эта процедура объединяет несколько отдельных объектов в один, что позволяет производить над ними групповые операции (масштабировать, перемешать, удалять и т.д.). Можно и, наоборот, разбивать объект, состоящий из нескольких, на самостоятельные объекты (разгруппировывать).

В графических редакторах предварительно необходимо выделить объекты, которые необходимо редактировать. Для выделения объектов в растровых графическом редакторе используют технологию выделения области. В выделенную область попадают все пиксели.

Выделение объектов в векторном редакторе осуществляется пообъектно. По этой причине векторные редакторы (редко) и векторные ГИС (чаще) называют объектными. Для выделения объекта можно с помощью указателя щелкнуть по любому объекту или с помощью других инструментов выделить площадь, в которой находятся объекты.

Редактирование позволяет вносить в объект изменения: стирать или изменять его части рисунка, изменять цвета или его копировать и вырезать.

Операцию изменения цвета в растровых редакторах можно осуществить с помощью меню *Палитра*, содержащего набор цветов, используемых при создании и рисовании объектов.

Графические редакторы позволяют добавлять в рисунок текст и осуществлять его форматирование.

Масштабирование или "зуммирование" применяют для детальной проработки фрагментов или просмотра изображения целиком.. Масштабирующие инструменты в графических редакторах позволяют увеличивать или уменьшать масштаб представления объекта на экране, но не влияет при этом на его реальные размеры.

В векторных редакторах существуют специальные процедуры типа построения фигур подобия, нахождение точки пересечения двух линий, редактирование по введенным координатам, обрезание по заданной границе и т.п.

## **6.6. Объектно-ориентированный подход в геоинформатике**

Как известно, разработка проекта (системы) на основе модернизации требует на порядок меньше затрат, чем разработка с проведением поиска и исследований. Это определяет одно из преимуществ объектно-ориентированного подхода - существенная эконо-

мия трудозатрат при проектировании систем, включающих известные решения и подходы.

Объект, с одной стороны, обладает определенными свойствами, которые характеризуют его состояние в данный момент времени, а, с другой стороны, над ним возможны операции, которые приводят к изменению этих свойств. Состояние объекта описывается стационарной или квазистационарной моделью. Изменение параметров объекта описывается динамической моделью.

Объектно-ориентированный подход (ООП) позволяет объединить обе модели, что является еще одним его достоинством. ООП строится на использовании классов объектов. Изменение свойств объекта в ООП не может быть произвольным. Оно возможно только через принадлежащие этому классу объектов методы. Методы «описывают» свойства объекта, что определяют как инкапсуляцию свойств в объект.

Инкапсуляцией называют объединение в объекте его свойств и возможных над ним операций (методов).

Инкапсуляция дает возможность уточнить понятие классов объектов. Объекты, инкапсулирующие одинаковый перечень свойств и операций, объединяют в *классы*. Поэтому отдельный объект является *экземпляром класса*. Экземпляры класса могут иметь отличающиеся значения свойств.

Например, все файлы в компьютерных системах обладают одинаковыми наборами свойств (имя, положение в файловой системе и др.) и операций (переименование, перемещение или копирование и др.) и образуют класс объектов *Файлы*. Каждый отдельный файл является экземпляром этого класса и имеет конкретные значения свойств (имя, местоположение и др.).

Таким образом, в объектно-ориентированном программировании центральное место занимают объекты, которые объединяют в одно целое (инкапсулируют) свойства объекта и возможные над ним операции (методы).

Если говорить языком баз данных, то объекты — это сущности. Объектом является, например, графический примитив *условный знак*. Свойства объекта, например, координаты, цвет, размер — его атрибуты. Методы объекта, это допустимый набор действий которые можно производить с объектом данного класса, например, переместить, генерализировать или детализировать.

В геоинформационных технологиях векторное изображение строится из графических примитивов (точка, линия, окружность и

т.д.), которые образуют классы геоинформационных объектов.

Один из классов графических примитивов образуют точечные объекты. *Точечный* объект обладает определенными свойствами (*Координаты, Цвет, Размер*), над которыми возможны соответствующие операции (*Перемещение, копирование, Изменение атрибута*).

Другой класс объектов в геоинформационных технологиях образуют линейные объекты. Они имеют дополнительные свойства и операции.

Классы объектов связаны между собой. Из более простых можно порождать более сложные. Например, из класса объектов *Точка* можно получить класс объектов *Окружность*, добавив новое свойство *Радиус* и операцию *Изменение радиуса*.

Все объекты класса *Окружность* наследуют свойства и операции класса *Точка*. Класс *Точка* называется *класс-родитель*, а класс *Окружность* — *класс-потомок*. В связи с возможностью использовать свойства порождающих объектов в порожденных вводят специальную характеристику наследование.

**Наследование** свойство объектно-ориентированного подхода, которое определяет отношение между классами объектов порождающих и классами объектов порожденных. Отношение наследования между классами можно выразить с помощью графа:

Часто встречается ситуация, когда над объектами различных классов можно совершать одинаковые операции. Так, в рассмотренном выше примере над объектами различных классов "Точка и" "Окружность" можно совершать одну и ту же операцию "перемещение".

Для большинства классов объектов в геоинформационных технологиях также характерен набор одних и тех же операций (переименование, перемещение, копирование, удаление и т.д.). Такое единообразие очень удобно для пользователя.

Механизмы реализации этих операций неодинаковы для различных классов. Эти операции будут выполняться различными программами. Таким образом реализуется еще одно свойство ООП - *полиморфизм*.

**Полиморфизм** - это возможность проведения одних и тех же операций над объектами, принадлежащими разным классам, при сохранении индивидуальных методов их реализации для каждого класса.

Полиморфизм в переводе с греческого означает «много форм». Одна и та же операция над объектами различных классов может вы-

полняться различными методами.

При наличии одинаковых наборов свойств и методов, объекты образуют *класс объектов*. Класс объектов может содержать множество различных объектов (*экземпляров класса*), каждый из которых имеет свои атрибуты имя. Объекты как правило образуют некоторую иерархию.

В геоинформатике используют три основных класса графических объектов: точечные, линейные и полигональные.

Объекты этих классов обладают также определенным набором методов, например: *открытие документа* (Open), *сохранение документа*, *построение буферной зоны* и т.д.

Для того, чтобы выполнить какую-либо операцию над объектом необходимо задать метод. Многие методы имеют аргументы, которые позволяют задать параметры выполняемых действий.

Обратиться к методу объекта можно с использованием стандартной для объектно-ориентированного подхода *точечной нотации*. Для того, чтобы определить для какого объекта вызывается метод, перед именем метода необходимо указать имя объекта, отделенное точкой.

**Синтаксис команды применения метода объекта:**

Объект.Метод arg1 —значение, arg2:=значение

Например, операция открытия в приложении документа map должна содержать не только название метода Open, но и указания пути к открываемому файлу (аргументу метода FileName необходимо присвоить конкретное значение).

Для того, чтобы изменить состояние объекта требуется задать новые значения его свойств. Обратиться к свойству объекта можно также с использованием *точечной нотации*. Указать изменяемые свойства объекта можно с помощью простой операции: перед именем свойства указывается имя объекта, отделенное точкой.

Для присваивания свойству конкретного значения используется знак равенства. Синтаксис установки значения свойства объекта:

Объект. Свойство = Значение Свойства

Одним из классов объектов является класс символов Characters(). Экземпляры класса нумеруются: Characters (1), Characters (2) и т.д.

Одним из примеров объектно-ориентированного подхода в программировании является язык визуального объектно-ориентированного программирования Visual Basic. Он основан на визуальном методе программирования, в частности, визуальное соз-

дание графического интерфейса приложения и объектный метод построения его программного кода.

В языке Visual Basic объекты характеризуются не только *свойствами* и *методами*, но и *событиями*. Событие представляет собой действие, распознаваемое объектом. Событие может создаваться пользователем (например, щелчок мышью или нажатие клавиши) или быть результатом воздействия других объектов приложения.

Для каждого события можно запрограммировать *отклик*, т.е. реакцию объекта на произошедшее событие. Например, если пользователь производит какое-либо воздействие на элемент графического интерфейса (нажимает *командную кнопку*), в качестве отклика выполняется некоторая последовательность действия (*событийная процедура*).

Имя процедуры включает в себя имя объекта и имя события.

Объект\_Событие()

В событийной процедуре может участвовать несколько объектов. Допустим, само событие происходит с первым объектом (Объект!), в результате второй (Объект2) изменяет значение своего свойства, а третий (Объект3) реализует какой-либо метод.

Каждая процедура представляет собой отдельный программный модуль, а синтаксис языка Visual Basic такой же, как языка VBA, поэтому запись событийной процедуры производится следующим образом:

Запись событийной процедуры на языке Visual Basic такова, что в начале и в конце ставятся ключевые слова Sub и End Sub.

Визуальное программирование позволяет создавать графический интерфейс разрабатываемых приложений на основе использования *управляющих элементов* (Controls), с которыми мы знакомимся при изучении графического интерфейса Windows.

Управляющие элементы чаще всего используются для получения от пользователя данных и вывода результатов работы приложения. Таким образом, управляющие элементы являются основой для построения пользовательского интерфейса приложения. 158

*Управляющие элементы* — это объекты, реагирующие на события, производимые пользователем, и являющиеся элементами графического интерфейса приложения.

Основными объектами, используемыми при визуальном программировании, являются формы (Forms). Форма представляет собой окно, на котором размещаются управляющие элементы. Форма обеспечивает создание графического интерфейса разрабатываемого

приложения.

*Форма* — это объект, представляющий собой окно на экране, в котором размещаются управляющие элементы.

Разрабатываемое на языке Visual Basic приложение называется *проектом*. Проект включает в себя не только форму с размещенными на ней управляющими элементами, но и программные модули событийных процедур, которые описывают поведение объектов приложения и взаимодействие объектов между собой. Проект — это разрабатываемое на языке Visual Basic, Delphi, Visual C++ приложение.

## 7. Хранение данных

Хранение данных в геоинформатике осуществляется в базах данных. Возможны три варианта организации баз данных.

1. База данных является подсистемой другой информационной системы. Такая организация имеет место в ГИС. В этом случае имеет место ассоциативная внутренняя связь информации базы данных с другими (например, графическими) данными.

2. База данных является внешней по отношению к информационной системе. В этом случае она соединяется с информационной системой или другой базой данных через интерфейс ODBC и возможна организация распределенной базы данных.

3. База данных подсоединяется к информационной или геоинформационной системе через Web-сервер. В этом случае возможна организация распределенной базы данных в режиме удаленного доступа.

Во всех вышеперечисленных ситуациях возникает необходимость проектирования базы данных, в качестве которой в настоящее время используют реляционную базу данных.

### 7.1. Проектирование реляционных баз данных

Геоинформационные системы могут содержать различные БД и подключаться к внешним базам данных.

Каждая БД объединяет данные, необходимые для решения одной или нескольких прикладных задач в рамках функционального назначения геоинформационной системы. Это могут быть кадастровые задачи, задачи автоматизированного картографирования, задачи геомаркетинга и др.

Базы данных подразделяют на прикладные и предметные. *Прикладные* обслуживают информационные приложения. *Предметные* охватывают некую предметную область, т.е. имеют более широкое назначение.

Предметные БД позволяют обеспечить поддержку любых текущих и будущих приложений, поскольку набор их элементов данных включает в себя наборы элементов данных прикладных БД. Вследствие этого предметные БД создают основу для обработки неформализованных, изменяющихся и неизвестных запросов и приложений (приложений, для которых невозможно заранее определить требования к данным).

Адаптивность позволяет создавать на основе предметных БД достаточно стабильные геоинформационные системы, в которых большинство изменений можно осуществить без вынужденного переписывания старых приложений.

В общем случае предметный подход проектирования БД используется для построения первоначальной информационной структуры, а прикладной – для ее совершенствования с целью повышения эффективности обработки данных.

При проектировании информационной системы содержащей БД проводят анализ ее целей и выявляют требования к ней.

Основная цель проектирования БД – это сокращение избыточности хранимых данных и оптимизация работы.

## **7.2. Универсальное отношение**

Проектирование базы данных начинается с анализа и выявления атрибутов и подбора данных. На основе анализа строится общая таблица всех атрибутов, которая может не являться отношением, так как ряд ее строк могут содержать не атомарные значения. Атомарными называют неделимые (простые) атрибуты.

Для придания таким таблицам формы отношения их реконструируют. Наиболее просто это сделать с помощью детализации сложных значений и вставки повторяющихся значений. Такое преобразование таблицы приводит к возникновению большого объема избыточных данных.

Тем не менее, полученная таблица представляет собой экземпляр корректного отношения. Ее называют универсальным отношением проектируемой БД. В одно универсальное отношение включаются все представляющие интерес атрибуты, и оно может содержать все данные, которые предполагается размещать в БД в будущем.

Таким образом, построение универсального отношения является первым шагом проектирования БД. Оно может использоваться в качестве отправной точки при проектировании БД.

## **7.3. Проблемы, требующие разбиения таблицы**

На практике при создании реальной БД приходится разбивать универсальное отношение (таблицу) на более мелкие. Это обусловлено тем, что при использовании универсального отношения возникает ряд проблем:

1. **Избыточность.** Данные ряда столбцов многократно повторяются. Могут повторяться и некоторые наборы данных. Эти повторения нежелательны.



**2. Потенциальная противоречивость (аномалии обновления).** Вследствие избыточности обновление данного в одной строке, оставляет его неизменным в других. Следовательно, при обновлениях необходимо просматривать всю таблицу для нахождения и изменения всех подходящих строк.

**3. Аномалии включения.** При включении в таблицу нового набора данных (часть из которых входит в другие наборы) возникает неоднозначность.

**4. Аномалии удаления.** Обратная проблема возникает при необходимости удаления атрибутов, входящих в удаляемый набор, но частично сохраняемых в других наборах.

Многие из перечисленных проблем этого примера исчезнут, если наборы данных выделить в отдельные таблицы и создать связующие таблицы

#### **7.4. Нормализации как метод построения нормальных форм**

Для того, чтобы разбить таблицу, называемую универсальным отношением на более мелкие необходимо какое-то правило или методика. Такой методикой является процесс нормализации.

**Нормализация** – процесс разбиения таблицы на более мелкие, обладающие лучшими свойствами при включении, изменении и удалении данных.

Конечная цель нормализации сводится к получению такого проекта базы данных, в котором каждый факт появляется лишь в одном месте, т.е. исключена избыточность информации. Это делается с целью экономии памяти и исключения возможной противоречивости хранимых данных.

Таблица, описывающая отношения, отвечает условию, в соответствии с которым в позиции на пересечении каждой строки и столбца таблицы всегда находится единственное атомарное значение, и, следовательно, никогда не может быть множества таких значений. Любая таблица, удовлетворяющая этому условию, называется нормализованной

Всякая нормализованная таблица автоматически считается таблицей в первой нормальной форме (сокращенно **1НФ**). В принципе "нормализованная" и "находящаяся в 1НФ" означают одно и то же. Однако на практике термин "нормализованная" часто используется в более узком смысле – "полностью нормализованная", который означает, что в проекте не нарушаются никакие условия нормализации.

В дополнение к 1НФ можно определить дальнейшие уровни нормализации – вторую нормальную форму (2НФ), третью нормальную форму (3НФ) и т.д.

По существу, таблица находится в 2НФ, если она находится в 1НФ и удовлетворяет некоторому дополнительному условию, суть которого будет рассмотрена ниже. Таблица находится в 3НФ, если она находится в 2НФ и, помимо этого, удовлетворяет еще другому дополнительному условию и т.д.

Таким образом, каждая нормальная форма является в некотором смысле более ограниченной, но и более желательной, чем предшествующая. Это связано с тем, что "(N+1)-я нормальная форма" не обладает некоторыми недостатками, свойственными "N-й нормальной форме". Общий смысл дополнительного условия, налагаемого на (N+1)-ю нормальную форму по отношению к N-й нормальной форме, состоит в исключении этих недостатков.

### **7.5. Функциональные и многозначные зависимости**

Теория нормализации основывается на наличии той или иной зависимости между полями таблицы. Определены два вида таких зависимостей: функциональные и многозначные.

**Функциональная зависимость** основана на зависимости полей внутри одной таблицы. Поле В таблицы функционально зависит от поля А той же таблицы в том и только в том случае, когда в любой заданный момент времени для каждого из различных значений поля А обязательно существует только одно из различных значений поля В. Отметим, что здесь допускается, что поля А и В могут быть составными.

**Полная функциональная зависимость.** Поле В находится в полной функциональной зависимости от составного поля А, если оно функционально зависит от А и не зависит функционально от любого подмножества поля А.

**Многозначная зависимость.** Поле А многозначно определяет поле В той же таблицы, если для каждого значения поля А существует хорошо определенное множество соответствующих значений В.

Для примера рассмотрим таблицу "Обучение" (рис. 7.1). В ней есть многозначная зависимость "Дисциплина-Преподаватель": дисциплина (в примере Информатика) может читаться несколькими преподавателями.

Таблица "Обучение "

Дисциплина	Преподаватель	Учебник
Информатика	Капранов П.А.	Макарова. Информатика
Информатика	Капранов П.А.	Уэйт М. и др. Язык Си
Информатика	Зайцев Г.Л.	Форсайт Р. Паскаль для всех
Информатика	Зайцев Г.Л.	Уэйт М. и др. Язык Си
Информатика	...Мишин И.В.	...Гооз Д. Информатика

Рис. 7.1. Иллюстрация многозначных зависимостей

Есть и другая многозначная зависимость "Дисциплина-Учебник": при изучении Информатики используются учебники "Информатика", "Паскаль для всех" и "Язык Си". При этом Преподаватель и Учебник не связаны функциональной зависимостью, что приводит к появлению избыточности (для добавление еще одного учебника придется ввести в таблицу две новых строки). Ситуация улучшается при замене этой таблицы на две: (Дисциплина-Преподаватель и Дисциплина-Учебник).

### 7.6. Нормальные формы

Приведем более строгое определение первой нормальной формы (1НФ), а также определения других нормальных форм.

Таблица находится в первой нормальной форме (1НФ) тогда и только тогда, когда ни одна из ее строк не содержит в любом своем поле более одного значения и ни одно из ее ключевых полей не пусто.

Таблица находится во второй нормальной форме (2НФ), если она удовлетворяет определению 1НФ и все ее поля, не входящие в первичный ключ, связаны полной функциональной зависимостью с первичным ключом.

Для упрощения нормализации подобных таблиц целесообразно использовать следующую рекомендацию.

**Рекомендация.** При проведении нормализации таблиц, в которые введены цифровые (или другие) заменители составных и (или) текстовых первичных и внешних ключей, следует хотя бы мысленно подменять их на исходные ключи, а после окончания нормализации снова восстанавливать.

Таблица находится в третьей нормальной форме (3НФ), если она удовлетворяет определению 2НФ и не одно из ее неключевых

полей не зависит функционально от любого другого неключевого поля.

Столкнувшись с несуразностями, которые могут возникать не только из-за введения кодированных первичных ключей, теоретики реляционных систем Кодд и Бойс обосновали и предложили более строгое определение для 3НФ, которое учитывает, что в таблице может быть несколько *возможных* ключей.

Таблица находится в нормальной форме Бойса-Кодда (НФБК), при условии, что любая функциональная зависимость между его полями сводится к полной функциональной зависимости от *возможного* ключа.

В следующих нормальных формах (4НФ и 5НФ) учитываются не только функциональные, но и многозначные зависимости между полями таблицы. Для их описания познакомимся с понятием полной декомпозиции таблицы.

**Полной декомпозицией** таблицы называют такую совокупность произвольного числа ее проекций, соединение которых полностью совпадает с содержимым таблицы.

Теперь можно дать определения высших нормальных форм. Начнем с определения последней из предложенных – 5НФ.

Таблица находится в пятой нормальной форме (5НФ) тогда и только тогда, когда в каждой ее полной декомпозиции все проекции содержат возможный ключ. Таблица, не имеющая ни одной полной декомпозиции, также находится в 5НФ.

Четвертая нормальная форма (4НФ) является частным случаем 5НФ, когда полная декомпозиция должна быть соединением ровно двух проекций. Весьма не просто подобрать реальную таблицу, которая находилась бы в 4НФ, но не была бы в 5НФ.

Так называемый, "чистый" проект БД ("каждый факт в одном месте") можно создать, используя методологию нормализации отношений.

## 7.7. Процедура нормализации

Как уже говорилось, нормализация – это разбиение таблицы на несколько, обладающих лучшими свойствами при обновлении, включении и удалении данных. Теперь можно дать и другое определение: нормализация – это процесс последовательной замены таблицы ее полными декомпозициями до тех пор, пока все они не будут находиться в 5НФ. На практике же достаточно привести таблицы к НФБК и с большей гарантией считать, что они находятся в 5НФ. Разумеется, этот факт нуждается в проверке, однако пока не существует

вует эффективного алгоритма такой проверки. Поэтому остановимся лишь на процедуре приведения таблиц к НФБК.

Эта процедура основывается на том, что единственными функциональными зависимостями в любой таблице должны быть зависимости вида  $K \rightarrow F$ , где  $K$  – первичный ключ, а  $F$  – некоторое другое поле. Заметим, что это следует из определения первичного ключа таблицы, в соответствии с которым  $K \rightarrow F$  всегда имеет место для всех полей данной таблицы. "Один факт в одном месте" говорит о том, что не имеют силы никакие другие функциональные зависимости. Цель нормализации состоит именно в том, чтобы избавиться от всех этих "других" функциональных зависимостей, т.е. таких, которые имеют иной вид, чем  $K \rightarrow F$ .

Если воспользоваться рекомендацией раздела 7.6. и заменить на время нормализации коды первичных (внешних) ключей на исходные ключи, то, по существу, следует рассмотреть лишь два случая:

1. Таблица имеет составной первичный ключ вида, скажем,  $(K_1, K_2)$ , и включает также поле  $F$ , которое функционально зависит от части этого ключа, например, от  $K_2$ , но не от полного ключа. В этом случае рекомендуется сформировать другую таблицу, содержащую  $K_2$  и  $F$  (первичный ключ –  $K_2$ ), и удалить  $F$  из первоначальной таблицы:

Заменить таблицу  $T(K_1, K_2, F)$ , первичный ключ  $(K_1, K_2)$ , функциональная зависимость  $K_2 \rightarrow F$   
на таблицу  $T_1(K_1, K_2)$ , первичный ключ  $(K_1, K_2)$ ,  
и на таблицу  $T_2(K_2, F)$ , первичный ключ  $K_2$ .

2. Таблица имеет первичный ключ  $K$ , не являющееся ключом поле  $F_1$ , которое функционально зависит от  $K$ . Кроме того таблица имеет другое неключевое поле  $F_2$ , которое функционально зависит от  $F_1$ . Решение здесь, по существу, то же самое, что и прежде – формируется другая таблица, содержащая  $F_1$  и  $F_2$ , с первичным ключом  $F_1$ , и  $F_2$  удаляется из первоначальной таблицы:

Имеется таблица  $T(K, F_1, F_2)$ , первичный ключ  $K$ , Функциональные зависимости  $K \rightarrow F_1$  и  $F_1 \rightarrow F_2$

Заменить ее на таблицу  $T_1(K, F_1)$ , первичный ключ  $K$ , и на таблицу  $T_2(F_1, F_2)$ , первичный ключ  $F_1$ .

Для любой заданной таблицы, повторяя применение двух рассмотренных правил, почти во всех практических ситуациях можно получить в конечном счете множество таблиц, которые находятся в "окончательной" нормальной форме и, таким образом, не содержат каких-либо функциональных зависимостей вида, отличного от  $K \rightarrow F$ .

Для выполнения этих операций необходимо первоначально иметь в качестве входных данных какие-либо "большие" таблицы (например, универсальные отношения). Но нормализация ничего не говорит о том, как получить эти большие таблицы. В следующей главе будет рассмотрена процедура получения таких исходных таблиц, а здесь приведем примеры нормализации.

## **8. Основы системного подхода при анализе и проектировании геоинформационных систем**

### **8.1. Системный анализ**

В настоящее время достаточно частым становится употребление термина "системный подход". В первую очередь его используют соискатели на ученую степень кандидата или доктора наук. Однако его применение не всегда соответствует содержанию этого понятия.

Как показывает анализ фонда диссертаций ВНИИЦ, именно методология системного анализа и сам анализ отсутствует в подавляющем большинстве работ, представляемых на соискание ученых степеней кандидатов и докторов наук [122].

Системный подход использует два основополагающих принципа:

- теория любых объектов или явлений основывается на некоем образе реальности, называемом моделью;
- формальные, инвариантные аспекты этой модели можно представить в виде структуры, построенной на основе неких логико-математических отношений.

Эти принципы дают основание выделить основные части системного подхода, применяемые при анализе информационных систем, включая ГИС: анализ структур, формализация, моделирование [5, 8, 74, 88, 98, ].

Сущность системного подхода заключается в проведении анализа сложной системы и представлении результатов этого анализа в виде некоего описания: графической схемы, аналитических выражений, совокупности моделей или статистических оценок. Следует отметить, что часто анализ производится не независимо в рамках перечисленных частей, а комплексно.

Основными целями системного анализа являются:

- определение структуры системы на уровне подсистем и

элементов;

- определение типов связей и информационных потоков между частями системы и внешней средой;
- определение функций частей системы;
- формализованное описание функций системы;
- построение параметрической модели системы и моделей ее частей;
- построение схемы системы;
- обобщенный анализ входных, выходных и промежуточных данных.

Системный анализ - это обобщенный анализ. В силу этого, а чаще из-за отсутствия необходимой информации, достижение всех перечисленных целей не всегда возможно. Поэтому при системном анализе информационной системы необходимо уточнять какие задачи ставились и какие результаты достигнуты.

Само по себе применение системного подхода не гарантирует новизну и положительный эффект исследования. Некорректное использование любой методологии не дает положительного результата и не служит характеристикой высокого уровня исследований.

На рис.8.1 показана укрупненная схема системного анализа информационных систем.

Она включает три основных группы процессов. Первая включает анализ структур системы. На основе этих процессов система разделяется на однородные по заданным признакам части.

Следующая группа процессов предназначена для создания формализованного описания найденных частей системы как целого механизма, отражающего связь между частями системы и функциональное назначение частей и системы в целом.

Еще одной группой процессов системного анализа являются процессы системного моделирования. Они разделяют параметры описания на две группы наблюдаемые и ненаблюдаемые. Это задает характер модельного эксперимента и технологии последующих экспериментальных работ.

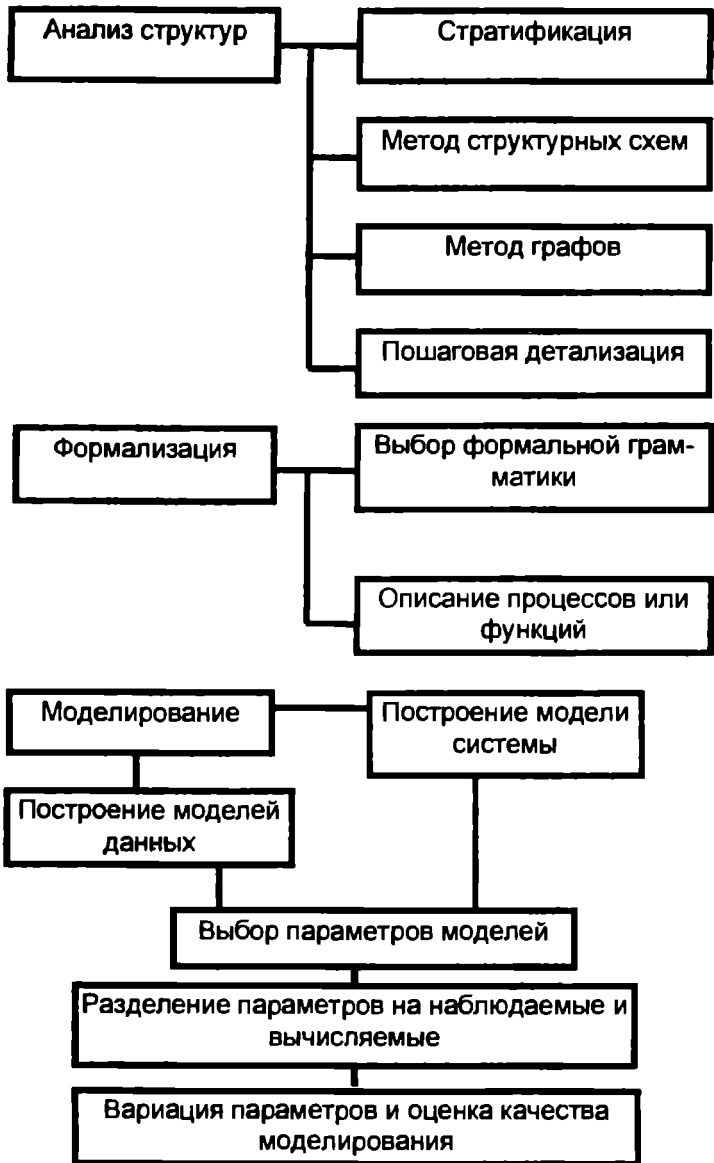


Рис.8.1. Основные этапы системного анализа



Процесс моделирования может видоизменяться. На его основе вырабатываются оценки качества моделирования и всего системного анализа.

Первый этап – анализ структур. На этом этапе при рассмотрении любой системы она представляет для исследователя некий “черный ящик”(рис. 8.2 а), структура которого неизвестна. Всякая система представляет собой сложное образование, имеющее части и элементы. Поэтому первым этапом анализа является выявление структуры системы, или используя терминологию тестирования преобразование “черного ящика” в “белый ящик” [122].

Для разбиения системы на крупные части (подсистемы) используют метод стратификации. Этот метод применяется в системном анализе, структурном анализе, археологии, но мало применяется в геоинформатике. Исследования этого метода связывают с М. Мессаровичем [64].

Суть его в разбиении множества исходных данных на уровни, которые называют стратами, отчего весь процесс называется стратификацией.

Образование уровней возможно в том случае, если множество характеристик, описывающих сложную систему, представимо в виде совокупности подмножеств, образующих декартово произведение [64, 98]. Примером такого произведения являются координаты точки в трехмерном пространстве. Другим примером являются векторные объекты, с которыми работает ГИС: точки, линейные, площадные объекты.

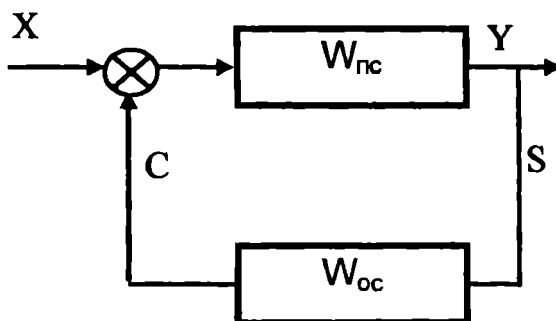
Стратификация превращает систему из неоднородного множества в совокупность иерархически взаимосвязанных однородных подмножеств. Она применима для анализа как систем, так и данных. Следует отметить, что стратификация эффективна для систем, которые могут быть описаны с помощью иерархических моделей.

Дальнейшая детализация структуры системы возможна с использованием метода структурных схем или метода графов.

Метод структурных схем мало используется в геоинформатике, его полное название “метод структурных схем и передаточных функций”. Он достаточно известен в теории автоматического регулирования [125] и широко применяется в структурном программировании. Он применим для анализа всех информационных систем (включая ГИС), предназначенных для решения задач управления и поддержки принятия решений.



а) "Черный ящик"



б) Структура с обратной связью

Рис. 8.2. Структурные схемы информационной системы

На рис. 8.2 приведена система регулирования в виде "черного ящика" свернутый вид рис. 8.2 а) и развернутый вид рис. 8.2 б).

Черный ящик характеризуется входным сигналом  $X$ , выходным -  $Y$ , и некой функцией преобразования этих сигналов  $\Phi$ . Эту функцию называют передаточной функцией.

$$Y(s) = \Phi(s) X(s)$$

Где  $(s)$  комплексная переменная в операторе преобразования Лапласа. При нулевых начальных условиях она аналогична оператору дифференцирования  $p=d/dt$ .

Метод структурных схем позволяет представить сложную систему в виде совокупности связанных подсистем ( на рисунке 8.2 б их две) и соответственно каждой подсистеме сопоставить свою пе-

редаточную функцию:  $W_{nc}$  - функция прямой связи;  $W_{oc}$  - функция обратной связи.

Развернутая схема на рис. 8.2 б) соответствует системе, имеющей обратную связь, и представляет собой типичную схему автоматизированного управления. Контрольный сигнал  $S$  поступает в регулируемую систему и в ней вырабатывается управляющий сигнал  $C$ .

Передаточная функция звена обратной связи  $\Phi_{oc}$  определится как

$$\Phi_{oc}(s) = \frac{W_{nc}(s)}{1 - W_{nc}(s) \cdot W_{oc}(s)}$$

Каждая функция в сложной системе может быть отображена своим узлом. Метод структурных схем удобен тем, что описывается с помощью операционного исчисления. Этим достигается согласование структурного и формального методов анализа.

Возможностей метода структурных схем может оказаться недостаточно, когда необходимо более детально исследовать процессы прохождения сигналов (данных) через каждое звено системы.

Такое представление дают сигнальные графы, разработанные С. Мезоном. Впоследствии теория графов была выделена для изучения как самостоятельная дисциплина [3]. Основными элементами графа являются узлы и дуги.

Для формализации описания графов Кирхгофом были предложены матрицы инциденций, т.е. матрицы соответствия дуг и вершин. В этих матрицах строки соответствуют вершинам, а столбцы дугам. Если дуга входит в узел, то на пересечении соответствующей строки и столбца элемент матрицы равен 1, если она выходит он равен -1, если дуга не касается, соответствующий элемент матрицы равен 0.

В такой формализации заложена топология, столь важная при анализе пространственных объектов. Матрица инциденций позволяет различать связанный и несвязанный графы. Несвязанному графу соответствует блочная квазидиагональная матрица.

Система состоит не только из подсистем, но и из элементов. Для выделения отдельных элементов применяют метод пошаговой детализации [98, 122].

Формализация описания системы включает два этапа: выбор грамматики (синтаксис и алфавит), собственно описание процессов в системе.

Моделирование включает построение моделей системы и мо-

делей данных [8, 10, 23, 24, 41, 60, 92, 126]. На первом этапе моделирования выбирают и обосновывают модель, определяют параметры необходимые для ее описания. После определения параметров оценивают какие из них являются наблюдаемыми и не наблюдаемыми.

**Наблюдаемые параметры** это те, которые можно измерить или получить в процессе эксперимента. Если параметры не наблюдаемы, то их необходимо заменять вычисляемыми параметрами, теми которые можно рассчитать по наблюдаемым параметрам. С этой целью могут быть введены дополнительные наблюдаемые параметры.

Таким образом, осуществляют основные этапы системного анализа. В целом следует отметить, что использование методов системного анализа повышает качество проекта системы, ее надежность, жизненный цикл и качество ее функционирования. Он позволяет выявлять ошибки на уровне обобщенного анализа и концептуальном уровне.

## **8.2. Системное проектирование в геоинформатике**

Системный подход к проектированию информационных систем или систем данных, применяемых в геоинформатике, к которым в первую очередь относят ГИС, использует ряд универсальных принципов.

Именно эта универсальность позволяет использовать системный подход как для проектирования функциональных систем (информацию в которых хранят или обрабатывают), так и при проектировании систем данных. По этой причине возможно применение термина **информационная среда** для обобщения не только функциональных систем (ГИС), но и для совокупностей данных (информационных моделей ГИС). В том случае, где возможно обобщение будет употребляться термин "информационная среда", в тех случаях, где речь идет только об информационной системе, будет употребляться соответствующее название.

При проектировании объектов информационной среды необходимо принимать во внимание совокупность системных принципов. В эту совокупность включают 15 основных принципов:

- 1 - системности;
- 2 - иерархичности;
- 3 - декомпозиции;
- 4 - интеграции;
- 5 – обратной связи;
- 6 – управляемости;

- 7 – контролируемости;
- 8 - согласованности;
- 9 - совместимости;
- 10 - реализуемости;
- 11 - единства системы и среды;
- 12 - преемственности;
- 13 - контринтуитивного проектирования;
- 14 - полиморфизации;
- 15 - синергетического эффекта

Кратко характеризуем каждый из названных принципов.

**Принцип системности** заключается в том, что любая информационная система содержит множество взаимосвязанных и взаимодействующих элементов, которые образуют взаимосвязанную систему. При этом исходная система является одновременно подсистемой более сложной системы.

Этот принцип требует, чтобы при анализе предметной области геоинформатики и формировании информационной среды были учтены не только существенные связи между элементами и уровнями ГИС или ГИС-технологий, но и связи с другими системами (внешние базы данных, системы и технологии сбора информации, технологии Интернет, пакеты САПР, ГИС с другими форматами данных, системы обработки изображений и др.), относящимися к окружающей среде.

Согласно данному принципу, при проектировании технологического процесса следует рассматривать конструкторско-технологические, информационные, экономические, социальные и экологические проблемы в комплексе друг с другом как единую систему проблем.

При реализации геоинформационных технологий следует учитывать, что изменение одного из его элементов может повлечь за собой последствия, порой и негативные, для других подсистем и самой технологий в целом.

**Принцип иерархичности** состоит в возможности выделения в сложной системе множества подсистем, находящихся на различных организационных уровнях, подчиненных друг другу по вертикали и имеющих собственные цели и функции, выполнение которых направлено на достижение общей цели системы.

Выделение подчиненных уровней в информационной системе необходимо для формирования комплексных критериев оценки эффективности функционирования системы.

Иерархичность достигается применением процедуры стратификации, описанной выше (см. также [5, 64, 98, 106]).

**Принцип декомпозиции** предполагает возможность детализации системы по тому или иному признаку на множества элементов и связей между ними и формирование для них собственных функций и целей из условия обеспечения исходных общесистемных функций и целей. Принцип используется для выделения элементного состава предметной области и связей между элементами при формировании информационного обеспечения системного проектирования.

Осуществление этого принципа возможно за счет применения инфологической модели "сущность-связь", описанной в разделе 3.4. данной книги.

При организации данных этот принцип применяют, например, для разбиения на слои исходной картографической информации.

**Принцип интеграции (композиции)** заключается в возможности объединения различными способами множества элементов (подсистем) с помощью множества связей в единую систему по единым правилам и в выявлении общесистемных свойств и функций.

В процессе проектирования данный принцип используется в задаче синтеза структуры информационной среды процесса и его обобщенных параметров из параметров подсистем и элементов подсистем.

**Принцип обратной связи** заключается в возможности обеспечения такого взаимодействия в системе, при котором выходные данные могут играть регулирующую роль в работе системы (рис.8.2.)

Этот принцип лежит в основе организации сетевых информационных потоков процесса проектирования.

Реализация этого принципа основана на применении метода структурного проектирования (раздел 8.1.).

**Принцип управляемости** заключается в том, что информационная среда не должна содержать элементов, подсистем и т.п., не реагирующих на управление.

Процесс обработки данных в геоинформатике должен быть управляемым. Для этого должны быть разработаны специальные интерфейсы, позволяющие эффективно принимать, передавать, обрабатывать информацию и вмешиваться в процесс обработки информации в необходимый момент.

При использовании ГИС в качестве системы управления пользователь должен иметь возможность вмешиваться в процесс принятия решений для использования собственных профессиональных знаний. Для этого проектируемый технологический процесс должен

иметь каналы управления.

**Принцип контролируемости** состоит в том, что информационная среда не должна содержать в своей структуре элементов, не контролируемых вышестоящим уровнем.

Для этого при проектировании должны быть предусмотрены контроль целостности данных, обнаружение и исправление ошибок. В проектируемом технологическом процессе должны быть включены операции контроля существенных параметров системы и качества данных в целях необходимого управления их состоянием.

**Принцип эффективности** заключается в необходимости обеспечения эффективности системы за счет согласованности ее элементов и оптимизации структуры системы применительно к технологическому процессу

При разработке или совершенствовании ГИС-технологий следует в первую очередь уделять внимание структурной оптимизации как основному резерву достижения функционального и экономического эффекта.

**Принцип согласованности** предполагает согласованность между собой всех элементов системы в целях достижения заданной эффективности.

Этот принцип реализуется за счет применения методов стандартизации. Этапы проектирования должны быть стандартизованы по программному, алгоритмическому и техническому обеспечению.

**Принцип реализуемости** предусматривает отсутствие в информационной среде элементов, не реализуемых имеющимися в распоряжении разработчика средствами техники и технологии.

Проектируемая процедура не должна содержать алгоритмов, не реализуемых с помощью располагаемых вычислительных средств. В разрабатываемом технологическом процессе должны отсутствовать операции и функции, которых не выполнимы с помощью выделенного объема ресурсов и располагаемых технических средств.

**Принцип единства системы и среды** предусматривает необходимость учета в проектируемой информационной системе параметров и свойств той среды, в которой предполагается ее функционирование.

**Принцип преемственности** предполагает возможность использования в новой версии проектируемой системы данных и методов более ранних версий. Проектирование должно осуществляться на основе широкой преемственности (стандартизированных проектных решений) с применением хорошо зарекомендовавших себя на практике типовых методик и средств формирования новых проек-

ных решений.

Реализация этого принципа осуществляется на основе объектно-ориентированного проектирования и программирования. Использование стандартов повышает надежность информационной среды.

**Принцип контринтуитивного проектирования** требует проектировать информационную среду, сочетая экспертный опыт с доказательными теоретическими решениями, представленными в аналитической форме. Реализацией этого принципа является использование экспертных систем.

**Принцип полиморфизации** заключается в возможности реализации сложных систем в многообразии форм (множестве морфизмов). Это условие является обязательным для интегрированной информационной среды.

Этот принцип обуславливает необходимость генерирования в процессе проектирования альтернативных вариантов технологий и параметров их компонентов. Другим его частным проявлением должно быть использование принципа многокритериальности, т.е. возможности оптимизации проектируемого технологического процесса по нескольким критериям. Необходимо, чтобы процедура проектирования была многоцелевой, ориентированной на разработку технологических процессов для различных типов производства.

На практике это означает, что одна и та же инструментальная ГИС может использоваться после ее настройки для разных целей. Например, она может использоваться как автоматизированная система картографирования, как система земельного кадастра, как система муниципального управления, как геомаркетинговая система, как система экологического мониторинга и т.д.

**Принцип синергетического эффекта** основан на получении эффекта, который присущ только системе как единому механизму. Синергетический эффект заключается в том, что, при компоновке или организации системы из совокупности компонент, эффект при взаимодействии компонент в рамках системы превосходит сумму эффектов от действия компонент независимо вне рамок системы.

Другими словами, объединение элементов в такую систему дает дополнительный эффект, который не имеет место при раздельном использовании частей или компонент системы.

Введение изложенных принципов в практику проектирования информационной среды будет способствовать созданию эффективных технологий в геоинформатике.

Эти принципы приемлемы при организации систем данных и при проектировании геоинформационных систем и технологий.



Универсальность системного подхода обеспечивает повышение эффективности работы комплекса: данные, информационная система, технология обработки или анализа данных. Особенно эффект системного анализа заметен при работе с большими объемами данных и сложными системами, что имеет место в геоинформатике.

## Прикладная геоинформатика

### 9. Сбор пространственно-временных данных в геоинформатике

#### 9.1. Исходные и унифицированные данные

Как уже говорилось, областью исследования геоинформатики являются объекты и явления земной поверхности [8, 26, 44–48, 51, 83–86, 108, 132, 133]. Для описания объектов реального мира нужна информация об их свойствах и характеристиках. Эту задачу решают технологии сбора информации.

По способу получения данные в геоинформатике подразделяют на *первичные* и *вторичные*.

**Первичные данные** - получают измерениями или наблюдениями непосредственно на исходном объекте [34]. Перечислим наиболее употребительные способы: аэрокосмическая съемка, выборочное обследование в полевых условиях, дистанционное зондирование, применение систем спутниковой навигации и др.

Первичные данные могут представлять собой некие наборы (например, записи в полевом журнале) или первичные модели (например, исходный снимок объекта).

**Вторичные данные** - это данные, которые получают на основе обработки первичных данных. Примером подобной обработки может служить решение прямой засечки по данным полевых журналов. Вторичные данные получают также из имеющихся моделей данных (например, сканированные изображения карт, снимков).

Различие между первичными вторичными данными несущественно для технологий обработки. Важным фактором является суммарная погрешность измерения координат точек исследуемых объектов. Поэтому в том случае, когда возникает проблема повышения точности обработки данных, следует проводить анализ и отдавать (при равных параметрах) предпочтение первичным источникам, поскольку они содержат меньше ошибок, вызванных методами измерения и обработки.

Получение информации о свойствах и характеристиках объектов осуществляют с помощью различных технологий. Каждая технология предназначена для сбора определенных типов данных.

Обрабатывать множество различных данных с помощью разных технологий неудобно и неэффективно. Значительно эффектив-

ней обрабатывать данные в рамках одной технологии. Такие технологии, позволяющие обрабатывать разнородные данные, называют *интегрированными* технологиями.

Для упрощения процесса обработки, хранения и обмена разнородные данные приводят к единому виду, который используется при обработке информации. Такие данные называют *унифицированными*.

Унификацией называется процедура сведения разнородных видов данных к единому виду. В ходе унификации данных осуществляется построение единой информационной модели, которая может быть использована в интегрированной технологии или интегрированной системе..

Класс, который образуют разнообразные исходные, нестандартизованные данные, называют *исходными данными*.

Внутренние, стандартизованные применительно к интегрированным технологиям данные, называют *унифицированными данными*.

Первый класс служит основой при сборе информации, второй - при обработке информации.

Поскольку исходные данные разнородны по стандартам, формам, представлению и т.д., они требуют некой предварительной обработки для их унификации. Этот этап обработки исходных данных называют *первичной обработкой* (рис.9.1). Его цель заключается в дополнении данных недостающей информацией, упрощении, исключение избыточности, анализе погрешностей, устранение или уменьшение погрешностей и т.п. Первичная обработка дополняет унификацию. Унификация по существу не изменяет информативность совокупности данных, а сводит их в информационную основу. При необходимости первичная обработка осуществляет анализ и изменяет информативность.

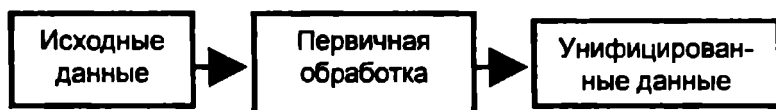


Рис.9.1. Преобразование исходных данных в унифицированные.

## 9.2. Основные характеристики пространственных объектов

Объекты реального мира, изучаемые в геоинформатике, имеют три основных типа характеристик: пространственные, временные и тематические.

**Пространственные характеристики** определяют положение объекта в заранее определенной системе координат. Традиционно связывают пространственное описание с координатными системами. Такой тип данных называют позиционным, поскольку он отражает ту часть информации об объектах, которая определяет (позиционирует) их положение на земной поверхности или в некоей заданной системе координат [5].

Основное требование к пространственным данным - точность. Это означает, что пространственные характеристики с требуемой точностью определяют положение объекта в системе координат и относительно других объектов.

**Временные характеристики** используются для того, чтобы указать время получения информации, накапливать временные ряды данных и иметь возможность исследовать протекание процессов и явлений с течением времени. Эти характеристики показывают зависимость изменения свойств объекта с течением времени.

Основное требование к временным данным - актуальность. Это означает, что данные можно использовать для обработки. Неактуальные данные это устаревшие данные, которые нельзя применять в новых изменившихся условиях.

**Тематические характеристики** описывают свойства объекта, не включенные в пространственные и временные. К ним относятся экономические, статистические, технические, организационные, управленческие и прочие виды данных. Основное требование к тематическим данным – полнота, которая определяется достаточностью для решения практических задач и нет необходимости проводить дополнительный сбор данных.

Помимо названных свойств пространственные, временные и тематические данные должны иметь ряд общих характеристик, например, надежность, достоверность и т.д. Однако главными требованиями к трем основным типам данных остаются вышеперечисленные.

Временные и особенно тематические данные фиксируют с помощью разных описательных характеристик, частично или полно-

стью формализованных. Поэтому эти типы данных определяют в виде кодов, дескрипторов, словарей и т.п.

Ряд геоинформационных технологий в качестве позиционных данных использует координатные данные, отчего этот класс данных называют *координатным*. В других технологиях этот класс данных определяют как *позиционный*. Эти термины следует считать синонимами в указанном смысле. Для определения параметров времени и тематической направленности применяют другой класс данных, называемый атрибутами (рис.9.2.).

На рисунке 9.2. показаны основные характеристики пространственных объектов, основные требования к этим характеристикам, основные классы данных, которые используются при организации информационной модели данных в геоинформатике.

Информационная модель - это объединяющая модель. Она включает всю информацию, получаемую при использовании разных технологий сбора. Следует иметь в виду, что информационная модель получается на основе первичной обработки исходных данных (рис.9.1.) и ее особенностью является то, что она включает новый класс данных, который при сборе информации в явном виде не присутствует. Это ассоциативные данные.

**Ассоциативные данные (или связи)** служат для связи позиционных данных с атрибутивными. Можно сказать, что ассоциативные данные связывают пространственные характеристики объектов с тематическими в единую систему. Ассоциативные данные превращают наборы независимых данных в систему связанных данных.

В теории баз данных такие данные называют метаданными [33, 56]. Они являются вспомогательными и незаметными («прозрачными») для пользователя. Из этого следует, что в процессе работы пользователь, обрабатывающий уже созданные наборы пространственно-временных данных не видит ассоциативные данные и не обязан их обрабатывать.

Ассоциативная связь позволяет при обработке пространственных данных получать соответствующие изменения в тематических характеристиках и наоборот. Ассоциативные данные позволяют после обработки тематической информации получать ее визуальное отображение с помощью пространственных данных, например в виде тематических карт.

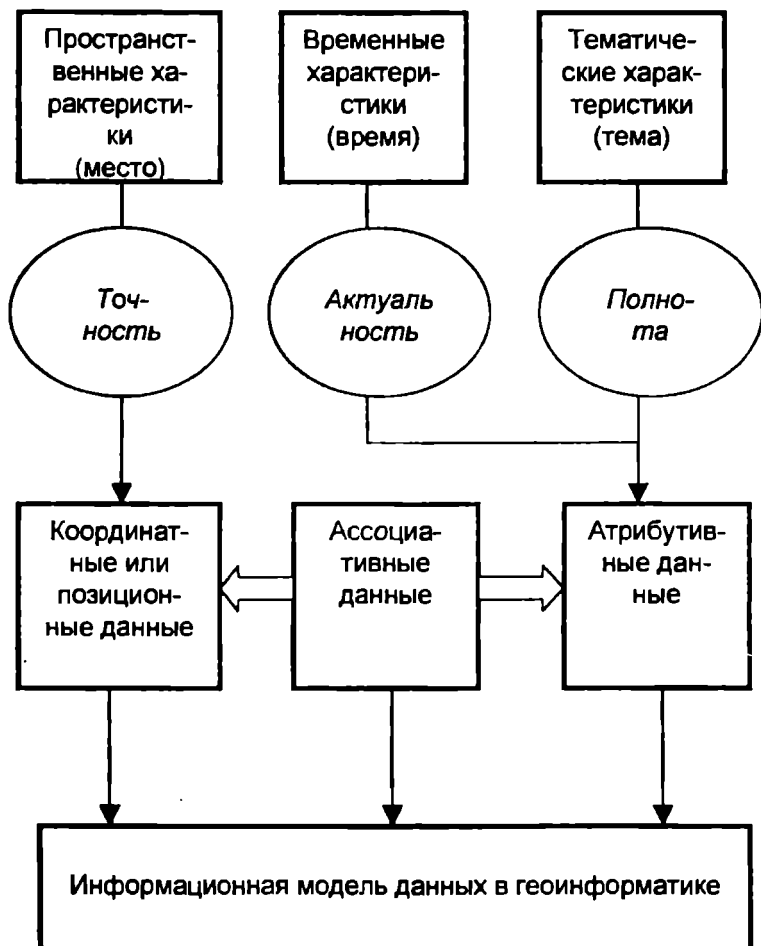


Рис.9.2. Основные характеристики объектов и организация информационной модели данных в геоинформатике

Примером использования ассоциированных данных может служить технология обработки данных в системе электронных таблиц Excel. Например, в этой системе можно создать следующую информационную совокупность:

- исходные табличные данные на одних листах;
- результаты обработки этих данных на других листах;
- деловая графика, по результатам обработки на третьих листах.

Если изменить исходные данные, то автоматически изменятся результаты обработки и графика, причем пользователь никаких действий для этого не сделает. Это пример ассоциативных связей.

Таким образом, информационная модель в геоинформатике содержит два основных класса данных: позиционный и атрибутивный, а также вспомогательный класс ассоциативных данных.

### 9.3. Форма и формат данных

Данные могут собираться с помощью различных технологий из различных источников. Общими характеристиками для всех данных будут формат, форма представления.

**Формат данных** – это способ их кодирования для обработки на компьютере. Обычно формат задается используемыми программными средствами. Преобразование данных одного формата в другой формат без изменения информативности осуществляется с помощью специальных программ- конвертеров.

**Форма представления данных** определяется способом их визуального представления. Она различна для координатных и атрибутивных данных.

Координатные данные могут иметь табличную, и графическую формы представления.

**Табличное представление** означает, что данные хранятся и представляются в виде таблиц.

**Графическое представление** означает, что данные можно представить в виде объектов компьютерной графики, которые с использованием технических средств можно преобразовать в бумажный оригинал.

Графические данные, служат основой визуального анализа и представления информации [2]. Их структура показана на рис. 9.3. Графические данные могут иметь векторное или растровое представление.

В свою очередь, векторные данные могут содержать информацию о топологии объектов или не содержать.

Топология определяет наличие связей между пространственными объектами и ее тип. Например, при слиянии двух рек имеет место один вид топологической связи, а при пересечении улиц другой. Кроме того, две дороги могут пересекаться, а могут проходить одна над другой.

При векторной нетопологической форме представления объект отвечает на вопрос: "*Где находится объект ?*", в то время как векторная топологическая форма представления отвечает на вопросы: "*Где находится объект и какие связи с другими объектами он имеет?*" (рис.9.3.). [56, 74].

При хранении в ГИС векторные данные должны обязательно иметь топологические характеристики. Определение топологии осуществляется либо непосредственно при сборе метрической информации, либо после сбора на основе дополнительного анализа и ввода этой информации.

Особо следует подчеркнуть, что различия между векторными и растровыми формами представления данных характерны именно для геоинформатики.

Растровые данные, как правило, представляют собой изображения и отображают *поля данных*, т.е. носят *полевой* характер.

Для векторных данных в геоинформатике характерно отображение геоинформационных *объектов*. По этой причине геоинформационные системы, использующие в качестве основного источника информации векторные данные, называют еще *объектными*.

Растровые данные получают чаще всего при сканировании. При сборе данных растровую форму переводят в векторную. Таким образом, преобразование растровой формы в векторную означает переход от полевого представления данных к объектному.

При переходе от растровой формы к векторной осуществляется селекция (выбор данных определенного типа) и существенное (в 100- 1000 раз) сжатие информации, но с сохранением информативности интересующих объектов.



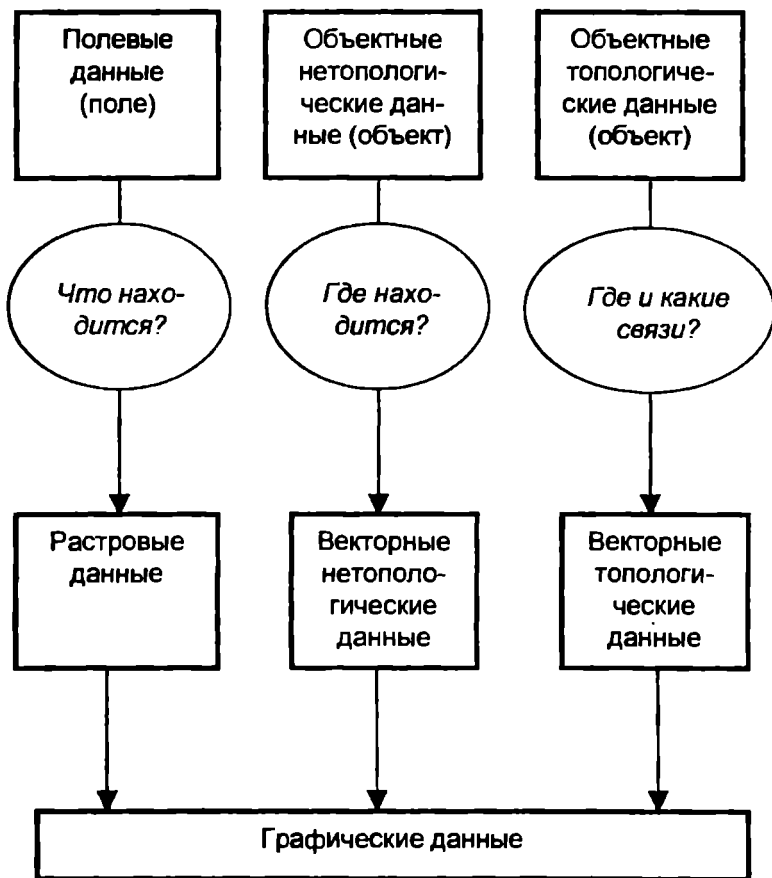


Рис. 9.3. Структура графических данных

Растрово-векторное преобразование (векторизация) применяется при интерпретации сканированных аэрокосмических изображений (выделения и оконтуривания на них однородных областей), в методах дигитализации цифровых растровых картографических изображений, при обработке данных, полученных с цифровых фотокамер и т.п. [10, 32, 70, 71]

Форма представления данных имеется как у позиционных так и у атрибутивных данных. Для атрибутивных данных она носит описательный характер и определяется типом выбранной структуры и модели хранения этих данных. В общем случае имеет место 5 форм представления атрибутивных данных:

- аналитическая (формулы, функции)
- графовая (структурированные схемы);
- графическая (графики, рисунки)
- табличная (совокупность таблиц);
- текстовая (тексты).

Наиболее часто атрибутивные данные имеют табличную форму представления. Атрибуты, соответствующие тематическим данным, определяют различные признаки объектов. Таблица, содержащая атрибуты объектов, называется таблицей атрибутов. В ней каждому объекту соответствует строка таблицы, каждому тематическому признаку – столбец таблицы.

Использование таблиц продиктовано тем, что таблица является основной информационной моделью в реляционных базах данных.

Временная форма может отражаться несколькими способами:

- фиксирование в одной таблице или в нескольких таблицах атрибутов данного объекта для различных временных этапов;
- указание временного периода существования объектов;
- соотнесение собираемой информации с определенными моментами времени;
- указание скорости движения объектов.

#### **9.4. Основные технологии сбора данных**

В геоинформатике используются данные, получаемые в различных технологиях [8, 12, 15, 32, 50, 58, 70-75, 77, 87, 89, 133]: в полевых условиях геодезическими (полевыми) методами; с помощью системы глобального позиционирования GPS;

фотограмметрическими методами (наземные и воздушные снимки);

с помощью средств и технологий дистанционного зондирования;

с карт ( географические, тематические, специальные и т.п.);

с помощью телевизионной видеосъемки;

по сети Internet;

из баз данных или из архивов;

из других ГИС;

с помощью средств мультимедиа;

На рисунке 9.4. показаны основные технологии сбора пространственно- временных данных в геоинформатике.

Материалы воздушной (аэросъемка или съемка с малых носителей), космической и наземной фотограмметрических съемок обрабатывают либо на специальных аналитических приборах, либо сканируют с последующей обработкой и выделением полезной информации.

Геодезические данные поступают на основе полевых измерений или данных получаемых с систем позиционирования (GPS). Возможно поступление данных по сети Интернет. Картографическая информация поступает на основе дигитализации или на основе цифровых картографических данных, уже хранящихся в других ГИС.

В настоящее время для ввода информации как дополнение к другим источникам информации все шире используют видеосъемку. Кроме того при вводе информации используют архивы ( каталоги координат), которые чаще всего хранятся в базах данных.

Обработка данных осуществляется в зависимости от технологии и инструментальной системы ГИС на разных вычислительных устройствах. Возможна обработка в режиме “клиент- сервер”, т.е. прибегают к услугам серверов и рабочих станций, но возможна обработка и в режимах Desktop GIS , т.е. на персональных компьютерах.

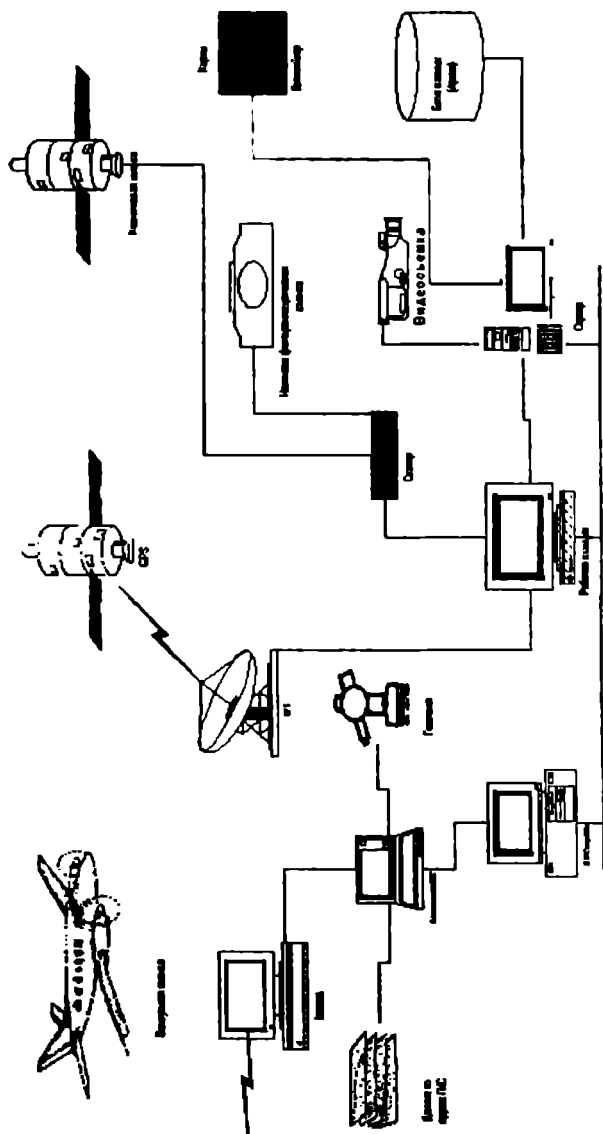


Рис. Сбор пространственно-временных данных для ГИС

### **9.5. Системы сбора полевой информации**

Полевая съемка непосредственно на местности определяет истинное горизонтальное и вертикальное положение объектов. При этом способе сбора информации осуществляют измерения: углов и расстояний.

Для съемки используют специальные геодезические инструменты (теодолиты, нивелиры и др.). По завершению полевых работ данные фиксируются либо в специальных полевых журналах, либо на устройствах автоматизированной регистрации в закодированном виде.

Эта информация, как было отмечено выше, называется исходной и требует первичной обработки и унификации.

При сборе полевой информации используют технологию, позволяющую уменьшать и даже исключать процесс предварительной (первичной) обработки полевых измерений перед вводом их в топографическую базу данных.

Такую автоматизированную технологию называют "сквозной". Она включает полностью автоматизированный процесс обработки геодезических измерений от полевых работ до данных цифровой модели (без записей в полевые журналы). Применение сквозной технологии возможно при использовании специальных геодезических приборов, оборудованных миникомпьютерами или вычислительными системами.

С помощью таких приборов (например, электронных тахеометров) наряду с процессом измерений проводят первичную обработку и унификацию данных.

### **9.6. Система глобального позиционирования GPS**

Одной из эффективных технологий сбора данных в геоинформатике является технология глобального позиционирования GPS. Определение данной системы Global Positioning System (GPS) переводится как глобальная система определения местоположения (позиционирования).

Система глобального позиционирования (GPS) является новой информационной технологией точного определения положения объектов на земной поверхности. Положение рассчитывается по сигналам, поступающим с серии искусственных спутников Земли (ИСЗ) NAVSTAR [57].

В основе системы заложено использование 23 спутников Земли, находящихся на околоземных орбитах на большой высоте.

Спутники расположены так, чтобы была возможность определения местоположения в любой точке Земли в течении 24 часов. Погрешность определения может составлять около 6 - 10 метров, а в дифференциальном режиме до одного сантиметра.

Основой определения местоположения служат специальные приемники, действие которых опирается на точные данные об орбитах спутников. Приемник с небольшой антенной способен определять свое положение в трехмерном пространстве с интервалом от 1 часа до менее 1 сек, в зависимости от используемого метода

Система глобального позиционирования (GPS) создана и применяется Вооруженными силами США, однако по специальному соглашению она доступна и для гражданских организаций.

Железнодорожные компании использовали GPS для первой в истории точной съемки железных дорог США и для слежения за движением поездов. С помощью GPS были уточнены высоты многих горных вершин, включая Монблан.

GPS обладают большим потенциалом и возможностями для интеграции с другими геоинформационными технологиями. В частности, для задач крупномасштабных съемок приходится выполнять большой объем геодезических измерений. Решение этой задачи эффективно осуществляется с применением системы GPS. Она позволяет решить две задачи:

- создание геодезического обоснования кадастровых планов;
- определение координат точек контуров (включающих границы земельных участков).

В абсолютном методе измерений координат используется только один приемник, в дифференциальном - два: один стационарный, а другой подвижный. Один приемник устанавливается неподвижно на точке с известными координатами, производит измерения дальности до всех видимых спутников, а также определяет постоянно меняющиеся характеристики тропосферы и ионосферы как сред распространения навигационных радиосигналов. По результатам обработки измерений стационарным приемником вычисляются дифференциальные поправки к дальностям и скоростям каждого спутника

Подвижный приемник последовательно устанавливается на точки с неизвестными координатами. Координаты точек местности могут определять как в реальном времени (on-line), так и в режиме последующей (off-line) обработки. Определение координат в реальном времени требует передачи поправок по радиоканалу от стацио-

нарного приемника на подвижный. Для этого выпускаются специальные приемники, которые называют базовыми станциями.

Анализ применения этой технологии показал, что она используется не только в геодезии и картографии, но и в землепользовании, экологии, наземной навигации, архитектуре и строительстве, геологии, региональном управлении, железнодорожном транспорте, дистанционном зондировании, образовании, метеорологии, демографии и здравоохранении и т.д.

Следует отметить эффективность применения такой системы для решения задач землеустройства. Практическое использование спутниковой навигационной системы выявляет ряд ее преимуществ при выполнении землеустроительных работ:

- точность определения координат пунктов выше, чем точность полученная средствами традиционной съемки;
- за счет отсутствия необходимости прокладки теодолитного хода временные затраты на определение координат точек меньше, чем при использовании традиционной наземной съемки;
- возможно проведение работ при отсутствии прямой видимости между измеряемыми пунктами;
- хорошее сочетание GPS-съемки с тахеометрической;
- возможность записи в процессе измерения координат во внутреннюю память приемника любой семантической информации;
- возможность экспорта результатов обработки в удобном для последующего использования ГИС-формате.

### **9.7. Сбор картографических данных**

Длительное время картографические данные служили основным источником данных для ГИС.

Карта как информационный носитель выполняет две функции:

- позиционную (дает информацию о точном расположении объекта, о его размерах);
- атрибутивную (информирует о типе, виде, классе объекта, показывает топологические свойства объектов, их отношений и т.п.).

Применяют два основных метода ввода картографической информации - сканерный и дигитайзерный. Соответственно, технологии сбора картографических данных называют сканерной и технологией дигитализации (цифрования). Охарактеризуем их более подробно.

**Технология дигитализации** (*digital* – «цифра» *англ.*) основана на дискретном преобразовании аналоговой информации, содержащейся на карте или ином графическом носителе информации (чертеж, фотоснимок) и создании новой дискретной (цифровой) карты. Вновь создаваемая карта может иметь другие характеристики по отношению к исходной. Эта карта создается таким образом, что при ее выводе на устройства печати будет создана аналоговая карта.

Таким образом, важным технологическим аспектом дигитализации является то, что в процессе дискретизации информативность создаваемой цифровой карты должна соответствовать информативности исходной аналоговой карте (разумеется в рамках решаемых задач). При этом следует отметить появление термина *цифровая карта* [109]. Он означает, что такая карта является *дискретным аналогом* бумажной карты и создается *для использования в компьютерных технологиях*, применяющих цифровые ЭВМ, а не аналоговые компьютеры, которые в настоящее время используют в редких случаях.

При данной технологии графический документ помещают на плоскую поверхность планшета и точку за точкой вводят информацию в компьютер. Для этой цели наводят на объекты карты специальное оптическое устройство (визир) и регистрируют каждую точку. Эта технология включает следующие этапы:

- закрепляют карту на графическом планшете;
- выбирают координаты контрольных точек (широта и долгота, плоские координаты, топографическая сетка) из каталога или другого источника;
- для каждого листа карты оцифровывают три контрольные точки или больше, если этого требует точность или выбираемая базовая картографическая проекция;
- выбирают контрольные точки так, чтобы их можно было легко найти по характерным объектам карты (перекрестки главных улиц, главные вершины, точки побережья),
- выполняют необходимые математические преобразования для определения коэффициентов связи между системой координат планшета и системой координат создаваемой карты,



- задают атрибуты создаваемых графических объектов (толщину и тип линии, цвет, изменение или постоянство толщины линии от начала к концу, слой и т.д.),
- пообъектно регистрируют точки на карте с помощью визира и координаты этих точек автоматически переводятся в координаты вновь создаваемой карты

Название технологии обусловлено в первую очередь тем, что непрерывная аналоговая информация с карты преобразуется в дискретную (цифровую) информацию. Дигитализация содержания карты осуществима двумя различными способами: ручным (одиночным) и автоматизированным (групповым).

При ручном способе оператор выбирает точки субъективно и фиксирует каждую точку. Из этого следует, что два оператора могут фиксировать одну и ту же линию по-разному

При групповом способе дигитализации точки фиксируются автоматически через заранее установленные интервалы времени или приращения координат. Для данного способа характерно образование большого количества точек, многие из которых могут оказаться избыточными.

Недостаток данной технологии в том, что размеры планшета должны быть сопоставимы с размерами источника. Кроме того, процесс ввода занимает продолжительное время.

Технология дигитализации имеет преимущество в том, что требует меньшего объема информации для файла, хранящего цифровую картографическую информацию и в том, что позволяет эффективно вводить небольшие фрагменты разных карт, соединяя их в общую цифровую модель. Она создает цифровую карту путем наращивания ее физического объема файла цифровой карты.

**Сканерная технология** основана на считывании карты, или другого бумажного графического документа с помощью специального считывающего устройства, называемого сканером. Она, в противоположность технологии дигитализации, создает цифровую карту путем уменьшения исходного большого физического объема файла, который образуется после сканирования.

Сканерная технология ввода картографической информации включает следующие этапы:

- карта или другой документ (снимок, чертеж) сканируются при помощи планшетного (редко) или при помощи барабанного (чаще) сканера, в результате получают растровый образ отсканированного документа;

- осуществляют геометрическую коррекцию полученного растрового изображения;
- выбирают координаты контрольных точек (широта и долгота, плоские координаты, топографическая сетка) из каталога или другого источника;
- в ручном режиме с помощью дисплея и мыши для каждого листа карты оцифровываются три контрольные точки или больше, если этого требует точность или выбираемая базовая картографическая проекция;
- рассчитывают формулы преобразований из системы координат сканированного изображения в систему координат карты или чертежа;
- осуществляют векторизацию (перевод из растровой формы представления в векторную) в ручном, полуавтоматическом ил автоматическом режиме;
- редактируют результаты векторизации в интерактивном режиме;
- записывают результаты обработки в базу данных.

Сканерная технология допускает возможность применения теории распознавания образов при автоматической векторизции. Это особенно выгодно при обработке карт. на которых имеется много стандартизованных образов, например, условные знаки [2]. Поэтому производительность сканерной технологии намного выше именно при вводе картографической информации, что в первую очередь определяется качеством автоматической векторизации.

### **9.8. Фотограмметрические методы сбора данных**

Фотограмметрическими называют методы сбора, использующие технологии получения и обработки различных фотоснимков [32, 50, 54, 70, 71, 74, 87, 89, 94].

Аэрофотоснимки и космические снимки как источники данных в геоинформатике играют существенную роль в выявлении различных природных факторов. Например, они позволяют идентифицировать антропогенную нарушенность природной среды, изучать морфометрию природных систем, проводить мониторинг экологического состояния территории.

Для уяснения использования фотограмметрических методов целесообразно сравнить информационные особенности снимка и карты.

*Позиционные характеристики.* Снимок более точен в передаче позиционных данных по сравнению с картой, поскольку при составлении карты объекты, изображенные на ней, могут искусственно разносить объекты для их лучшего показа.

*Соотношение полезной и избыточной информации.* Карта содержит необходимую информацию и не содержит ненужной, в то время как на снимке имеется много избыточной информации.

*Временной фактор.* Процесс подготовки, составления и издания карт методами обычной полиграфии может составлять несколько лет, что снижает актуальность информации представленной на карте. Снимки более оперативный источник данных, так как процесс их получения на один два порядка менее длительный, чем процесс издания карт.

*Тематический аспект.* Карта тематически более насыщена классифицированной тематической информацией. Традиционные фотограмметрические технологии ориентированы в основном на получение координат, (позиционных данных). При использовании фотограмметрических технологий тематическая информация должна вводиться дополнительно с помощью других технологий или дополнительных методов.

Для получения тематической информации по снимку его дешифрируют, т.е. осуществляют анализ и интерпретацию изображения. Однако для автоматизированной обработки дешифрированные объекты должны быть закодированы.

Таким образом, снимки содержат часть информации отображаемой на картах.

Снимки служат хорошим средством оперативного контроля и мониторинга. В отличие от карт они позволяют оперативно выявлять новые явления и процессы их развития [32, 70, 87, 89].

Снимки могут быть получены посредством космической съемки (с искусственных спутников Земли), аэрофотосъемки (со специально оборудованных самолетов, имеющих аэрофотокамеры), съемки с малых носителей (вертолеты, мотодельтапланы, авиамодели), наземной съемки (фототеодолиты, камеры для наземной съемки, любительские фотокамеры).

Фотограмметрические методы сбора данных дополняют другие источники информации в геоинформатике (рис.9.4) [32, 70, 74, 87, 89].

В отличие от технологий сбора картографических данных, позволяющих непосредственно создавать слои электронных карт в

ГИС, фотоснимки требуют предварительного применения процедур фотограмметрической обработки, в результате которых получают трансформированные изображения, цифровые модели и т.п.

Именно результат фотограмметрической обработки снимков является основой для создания картографических слоев электронных карт.

В зависимости от качества и условий съемки фотограмметрическая обработка может быть реализована разными программными и техническими средствами.

Например, трансформирование и изменение масштаба снимков на основе аффинных или проективных преобразований с дополнительным преобразованием координат снимка в заданную картографическую проекцию можно осуществить стандартными средствами ГИС.

При получении снимков в разных условиях или с различными параметрами съемки необходимо метрическое согласование информации получаемой по разным снимкам. Поэтому необходимо добиться геометрически правильного изображения всей исследуемой территории, что обеспечит хорошую сопоставимость данных.

Для обеспечения снимков необходимым количеством точек может быть используется аналитическая фототриангуляция или в простых случаях цифровая плановая фототриангуляция. Применение плановой фототриангуляции позволяет упростить технологию проведения работ, по сравнению с классической аналитической фототриангуляцией.

В настоящее время интенсивно развивается цифровая фотограмметрия. В цифровой фотограмметрии снимки получают в цифровой (дискретной) форме на машинных носителях информации. Такие снимки называют цифровыми снимками.

Цифровая фотограмметрия основана на аналитической компьютерной обработке цифровых снимков. Цифровой снимок получает-ся двумя путями:

- обычный снимок сканируется и переводится в растровую форму;
- применяют цифровые камеры, в которых вместо фото-пленки используют специальные элементы, вследствие чего снимок сразу получается в дискретной форме и может вводиться в компьютер.

Совершенствование цифровых фотокамер идет по двум направлениям:

- уменьшение габаритов съемочных камер до размеров любительских фотокамер с сохранением высокой точности съемки;
- разработка цифровых фотокамер, позволяющих без использования сканеров, получать высокоточное растровое изображение на запоминающее устройство и вводить его сразу в компьютер.

В заключении следует отметить, что фотограмметрические методы сбора информации основаны на получении снимков в отраженном свете. По этой причине фотограмметрическое оборудование называют **пассивными сенсорами**.

Использование фотограмметрических методов требует наличия определенной освещенности объекта и определенного времени экспозиции. Эти требования ограничивают применимость фотограмметрических методов.

### **9.9. Применение методов сбора и обработки ДДЗ в геоинформатике**

В настоящее время для съемок из космоса используются различные виды технических средств, которые включают [32, 70, 74] :

- системы с традиционной фотографической регистрацией изображения (фотопленка, снятая в космосе, на Земле проявляется и сканируется на прецизионных прожекторных сканерах для использования в геоинформационных технологиях);
- сканирующие системы реального времени, оперативно передающие информацию на Землю по каналам связи;
- радиолокационные (радарные) съемочные системы, в том числе так называемые радары с синтетической апертурой (РСА).
- фотографические системы КАТЭ-200, КФА-1000, КФА-3000 и МК-4/
- системы получения инфракрасных (тепловых) снимков.

Получение данных дистанционного зондирования для исследования природных ресурсов Земли является целевым назначением отечественных искусственных спутников Земли (ИСЗ) “Ресурс-Ф1”, “Ресурс-Ф2” и “Ресурс-Ф3” и др.

На борту спутников “Ресурс-Ф1” с 1974 года установлена 3-х канальная многозональная фотографическая камера КАТЭ-200 и две

смонтированные камеры КФА-1000, ведущие спектрзональную съемку с высоты 270 км.

С 1978 года ведется фотограмметрическая съемка одноканальной панхроматической камерой КФА-3000 с борта ИСЗ "Ресурс-ФЗ" (высота орбиты - 270 км).

Использование ИСЗ "Ресурс-ФЗ" (средняя высота орбиты также 270 км) с установленной на его борту многозональной фотокамерой МК-4 началось в 1987 году. В процессе съемки из космоса применяют разные типы аппаратуры, которая определяет точность и применимость ДДЗ.

### 9.9.1. Многозональная съемка

Говоря о фотограмметрических методах дистанционного зондирования, мы употребили термин «многозональная» или «зональная съемка». Что кроется за этими наименованиями?

В обычной фотограмметрической съемке изображение объекта фиксируется на один снимок (цветной или черно белый).

Реальное изображение, как известно из курса физики и оптики, имеет спектр. Это означает, что в разных оптических (видимых человеческому глазу) диапазонах имеется разная интенсивность отраженного света. При фотографировании на один снимок все эти излучения в разных диапазонах накладываются друг на друга, а на снимке получается суммарная интенсивность разных спектральных диапазонов.

Такое наложение спектров может мешать анализу изображения. *Многозональная съемка* основана на разделении всего спектрального диапазона на диапазоны, в которых раздельно фиксируют изображение. Таким образом, вместо одного снимка получают несколько, каждый из которых содержит изображение заданного спектрального диапазона, чем облегчает анализ и интерпретацию изображения. При использовании цифровых методов обработки изображений из спектрзонального изображения легко получить обычное. Однако спектрзональные изображения значительно проще анализировать, особенно с применением методов автоматической обработки данных.

### 9.9.2. Инфракрасная съемка

Выше мы говорили, что фотоснимок получается в диапазоне, видимом человеческим глазом. Это достигается специальным изготовлением фотоматериалов, фиксирующих излучение видимого диапазона.

Современные технологии позволяют получать такие фотоматериалы, которые фиксируют и переводят в видимый диапазон излучения не видимые человеческим глазом. К таким излучениям относится инфракрасная или тепловая съемка.

Инфракрасная или тепловая съемка основана на получении снимков, фиксирующих излучение в тепловом диапазоне, который человеческий глаз не фиксирует. В результате этой съемки невидимые излучения становятся доступными для визуального анализа.

Инфракрасная съемка, как и фотограмметрическая, основана на фиксации отраженного излучения. Это дает основание относить аппаратуру инфракрасной съемки к пассивным сенсорам.

Инфракрасная съемка дополняет фотограмметрическую, поскольку фиксирует явления и объекты, не получаемые на обычных снимках. Она является незаменимой при мониторинге пожаров и разных физических явлений, связанных с выделением тепловой энергии, например мониторинге тепловых или атомных электростанций.

В то же время следует отметить недостаток инфракрасных съемок - более низкая метрическая точность по сравнению с фотограмметрической.

### 9.9.3. Радиолокационная съемка

Радиолокационные (радарные) снимки отличаются от всех остальных данных дистанционного зондирования, как по методам получения и обработки, характеристикам, так и по специфике их использования.

Радиолокационные снимки получают на основе облучения объекта наблюдения активной станцией (радаром) и фиксации отражения этого излучения.

В отличие от других космических съемочных систем (пассивных сенсоров), регистрирующих отраженное оптическое излучение или тепловое излучение, радар - активный сенсор.

Радар не нуждается во внешних источниках освещения (облучения). Он как бы сам "освещает" исследуемую территорию, что определяет основное достоинство радарных снимков - независимость их получения от времени суток и года, погодных условий.

Облачный покров совершенно прозрачен для радара. - Радарные снимки незаменимы также для съемки приполярных районов в условиях полярной ночи и низкого стояния солнца!

На радиолокационных изображениях достаточно четко распознаются формы рельефа и характер поверхности, что обуславливает

их широкое применение в геологии и при выявлении участков волнения акватории

Радарное излучение способно проникать на некоторую глубину в почву и глубина этого “проникновения” зависит от влажности почвы. Это обеспечивает возможность в ряде случаев получать данные о рельефе коренных пород под рыхлыми, особенно песчаными отложениями или о следах искусственных сооружений прошлых эпох.

На радиолокационных снимках ясно выделяются техногенные объекты, например, хорошо отражающие радиоволны металлические конструкции.

В современных методах получения радарных снимков используются радары с синтетической (синтезированной) апертурой (РСА), радиолокация проводится в нескольких частотных диапазонах (на различных длинах волн).

Первичные данные РСА представляют собой не привычное изображение, а так называемую “голограмму” по которой синтезируется сам снимок. Этот синтез проводится по участкам (частям), поэтому потом требуется восстановление единого целого изображения.

Далее для обеспечения равномерности снимка осуществляется геометрическая коррекция строк изображения по дальности и по азимуту. Радиометрическая коррекция включает в себя коррекцию яркости по строке дальности радиолокационного изображения с учётом неравномерности коэффициента усиления антенны и коррекцию яркости элементов изображения с учетом дальности распространения радиоволн.

Таким образом, радиолокационные изображения обладают специфическими геометрическими и яркостными характеристиками, поэтому для их эффективного применения необходимо использовать специальные математические методы и программные средства.

В целом, область применения радарных съемок включает:

- - геологию;
- - археологию;
- - лесное хозяйство;
- - экологический мониторинг (особенно распространение пятен нефтяного загрязнения на акватории);
- - мониторинг влажности почв и состояния посевов;
- - мониторинг ледовой обстановки и проводка судов в Арктике;
- - чрезвычайные ситуации (в частности, вызванные наводнениями);



- -мониторинг береговых геоморфологических процессов;
- -военное дело и разведку.

Следует отметить недостаток радиолокационных съемок - более низкая метрическая точность по сравнению с фотограмметрической. Другим недостатком является зависимость от отражающих свойств поверхности (например влажности). При изменении этих свойств с течением времени радиолокационные снимки одной и той же поверхности, сделанные в течении разных периодов наблюдений, могут существенно отличаться.

## 10. Координатные и атрибутивные данные и модели.

Выше было установлено, что основой интеграции в геоинформатике служат координатные данные. Они используются для определения места и образуют один класс данных. Для определения параметров времени и тематической направленности существует другой класс данных - а т р и б у т ы (см. п.9.2).

Координатные *данные* служат основой организации класса координатных *моделей*.

### 10.1. Координатные модели

Одно из основных назначений координатных моделей отображать пространственные свойства объектов. Пространственные объекты имеют графическую форму представления посредством которой решаются две задачи:

- показывается взаимное расположение объектов и связи между ними (топология);
- дается количественная оценка геометрических характеристик объектов и их положения в выбранной системе координат (метрика).

Графическая информация, с помощью которой отображают отражающие трехмерную реальность пространственные объекты, включает:

наборы точечных (локализованных) объектов;  
 разнообразных линий и линейных (сетевых) объектов;  
 совокупности контуров и площадных объектов.

Это группы координатных данных служат основой для организации класса координатных моделей и трех групп координатных моделей в геоинформатике.

Таким образом, в геоинформатике приняты следующие основные типы координатных моделей (рис. 10.1):

точка (узлы, вершины);

линия (незамкнутая);

контур (замкнутая линия);

полигон (район, ареал) - группы примыкающих друг к другу замкнутых участков.

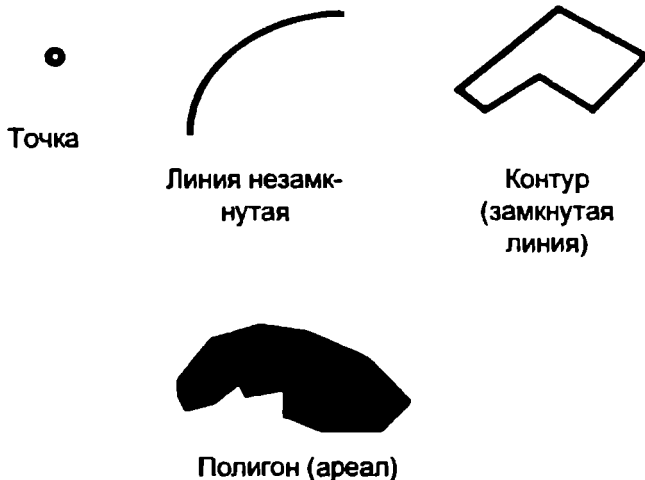


Рис.10.1. Типы координатных моделей

В некоторых системах в описание основных типов моделей включают понятие *пространственная сеть*, которая является развитием типа данных *район*. Контурные и линии часто объединяют общим термином - "линейные объекты". Таким образом, число основных типов координатных моделей меняется от трех до пяти в зависимости от используемой ГИС.

### 10.1.1. Точечные модели

Точечные координатные модели - это модели, отображающие пространственные объекты, локализованные в небольшой части пространства.

Точечные модели содержат два основных свойства и отвечают на два вопроса.

1. Они показывают место, *где* находится объект, например, в системе географических координат.
2. Они дают краткую характеристику объекту и показывают, *какой* это объект.

Данные модели не определяют размеры объекта. К этому типу моделей относятся не только точки, но и все точечные условные знаки. Выбор объектов, представляемых в виде точек, зависит от масштаба карты или исследования. Так, на крупномасштабной карте точечными моделями показываются отдельные строения, а на мелкомасштабной карте - города.

Особенность точечных моделей пространственных объектов состоит в том, что они могут храниться в двух видах:

- в виде графических файлов, как другие пространственные объекты;
- в виде таблиц, как атрибуты.

Последнее обусловлено тем, что точечные модели всегда характеризуются только двумя координатами. В силу этого информацию о точечных моделях можно всегда хранить в виде таблицы, содержащей одинаковое число столбцов. Кроме координат в этой таблице могут быть идентификационные номера, тематические характеристики и т.д. В таких таблицах каждая строка соответствует одной точечной модели - в ней собрана вся информация. Каждый столбец - это признак, содержащий типизированные данные (координаты или атрибуты). Каждая точка независима от всех остальных точек, представленных отдельными строками.

Точечные модели не имеют топологических характеристик.

### 10.1.2. Линейные модели

**Линейные координатные модели** - это модели, отображающие пространственные протяженные объекты с помощью линий или сетей. Эту модель образует совокупность связанных линий, что и определяет ее название.

Линейные модели отвечают на следующие вопросы.

1. Они показывают место, *где* находится объект, например, в системе географических координат.

2. Они показывают, *какой* это объект и дают ему краткую характеристику.

3. Они показывают, *каковы линейные размеры* объекта.

4. Они показывают, *в какой связи находится* объект по отношению к другим.

Таким образом, для линейных моделей в отличие от точечных характерно присутствие топологических признаков. Любая линейная модель включает узлы (вершин) и линии (дуг). Для каждого узла существует топологическая характеристика – валентность узла, которая определяется количеством отрезков линий, пересекающихся в нем.

Например, концы обособленных линий одновалентны. Для уличных сетей (пересечения типа “крест”) наиболее характерны четырехвалентные узлы. В гидрологии чаще всего встречаются трехвалентные узлы.

В иерархической модели каждая пара узлов имеет лишь одно соединение, не допускаются петли и замкнутые контуры. Эта модель описывает часть речных сетей (без островов). Наличие островов требует применения сетевой модели.

Линейные объекты, как и точечные, имеют свои атрибуты, причем разные для дуг (звеньев) и узлов.

Для точечных объектов необходимо указать одно значение координат, для линейных - как минимум два (для начальной и конечной точек). В более общем случае количество точек в линейной модели произвольно.

### 10.1.3. Площадные модели

**Площадные координатные модели** - это модели, отображающие пространственные объекты с замкнутыми границами и наличием каких-то свойств внутри этих границ. Эти объекты называют также ареалами, или полигонами.

Площадные модели отвечают на следующие вопросы.

1. Они показывают место, *где* находится объект, например, в системе географических координат.

2. Они показывают, *какой* это объект и дают краткую характеристику объекту.

3. Они показывают, *каковы линейные размеры* объекта.

4. Они показывают, *каковы площадные размеры* объекта.

5. Они показывают, *каковы свойства границ* объекта.

6. Они показывают, *каковы свойства объекта* внутри границ.

7. Они показывают, *в какой связи находится* объект по отношению к другим объектам

Таким образом, можно утверждать, что площадные модели отвечают на большее число вопросов и являются наиболее информативными из координатных моделей. Площадные координатные модели, как и линейные, могут иметь топологические характеристики.

В площадных координатных моделях границы могут определяться свойством или явлением, а также независимо от явления (затем перечисляются значения атрибутов).

Границы могут соответствовать границам пространственного объекта (государственная граница), явления (зона наводнения или пожара) или устанавливаться искусственно (например, для микрорайонов).

## 10.2 Атрибутивные модели

Координатных моделей недостаточно для описания тематической и графической информации. В последнем случае речь идет о неких связях между графическими объектами, которые в силу масштаба или большой плотности графических элементов не могут быть показаны в графическом виде.

К атрибутивным данным относятся: символы, названия, статистическая информация, код объекта, графические признаки (цвет, рисунок заполнения контуров) и т.д.

Как видим, все дополнительные описательные характеристики входят в атрибутивные данные. Таблица является *основной формой* хранения этих данных. Поэтому *основной моделью* атрибутивных данных является реляционная модель. Таблица, содержащая атрибуты объектов, называется *таблицей атрибутов*.

Каждому объекту соответствует строка таблицы, каждому тематическому признаку - столбец таблицы. Клетка таблицы отражает значение определенного признака для определенного объекта.

Для создания атрибутивных моделей данные должны быть типизированы, разделены на группы и ассоциированы (связаны) с пространственными объектами.

Особенность организации атрибутов в том, что одна тематическая характеристика может относиться к нескольким пространственным объектам. Эта особенность учитывается с помощью вспомогательных (ассоциативных) данных, которые называют также метаданными.

Выше отмечалось, что атрибутивные данные описывают тематические и временные характеристики. Временная характеристика может отражаться несколькими способами:

- путем указания временного периода существования объектов;
- путем соотнесения информации с определенными моментами времени;

- путем указания скорости движения объектов.

В зависимости от способа отражения временной характеристики она может размещаться в одной таблице или в нескольких таблицах атрибутов данного объекта для различных временных этапов.

Применение атрибутов позволяет осуществлять хранение и анализ пространственных объектов с использованием баз данных.

Таким образом, атрибутивное описание дополняет координатное, совместно с ним создает полное описание моделей ГИС и решает задачи типизации исходных данных, что упрощает процессы классификации и обработки.

### **10.3. Точность координатных и атрибутивных данных**

Использование любой информации допустимо, если она удовлетворяет определенным критериям и стандартам. Одним из критериев применимости пространственно-временных данных в геоинформатике является точность[30].

Точность определяют как - близость результатов, расчетов или оценок к истинным значениям (или значениям, принятым за истинные).

Например, точность горизонтали, полученной после дигитализации или сканирования, можно оценить сравнением ее с горизонталью на исходной карте.

Перечислим основные показатели точности в геоинформатике: точность вычисления, точность измерения, точность представления, позиционная точность, точность атрибутов.

Первые три характеристики связаны с числовыми преобразованиями и могут характеризовать любую информационную систему. Последние две являются характерными для геоинформатики.

**Точность вычисления** определяется количеством значимых цифр после запятой.

**Точность измерений** определяется количеством значимых цифр при измерениях.

**Точность представления** определяется количеством разрядов, описывающих координатные данные.

Термин "значимые" цифры означает, что данная цифра несет полезную информацию. Например, при вычислении можно рассчитать число с 20 знаками после запятой. Но если алгоритм вычисления создает ошибку в третьем знаке после запятой, то писать все знаки после второго не имеет смысла.

Точность вычисления обычно выше, чем точность измерений. Точность входит в комплекс данных, определяющий важный показатель - качество данных.

**Позиционная точность** выражается степенью отклонения данных о местоположении от истинного положения объекта на местности. Обычно точность карт приблизительно определяется толщиной линии или 0,4 мм. Это соответствует 10 м в масштабе 1:25 000.

Для проверки позиционной точности используют независимые более точные источники, например карту более крупного масштаба, систему глобального позиционирования (GPS) и др.

**Точность атрибутов** определяется близостью значений атрибута к его истинной величине. Атрибуты могут со временем меняться довольно часто по сравнению с координатными данными.

#### **10.4. Векторные и растровые модели**

Визуальное представление данных в геоинформатике основано на компьютерной графике. Основу компьютерной графики образуют векторные и растровые модели.

В общем случае модели пространственных (координатных) данных имеют векторное или растровое (ячейчатое) представление, содержат или не содержат топологические характеристики. Исходя из этого графические модели классифицируют по трем типам:

- растровая модель;
- векторная нетопологическая модель;

- векторная топологическая модель.

Все эти модели взаимно преобразуемы. Основу такой классификации составляет единица пространства, содержащая представления площадей линий и точек.

#### 10.4.1. Векторные модели

В отличие от занимающих все пространство растровых моделей векторные модели данных строятся на линиях, занимающих часть пространства. Это определяет основное преимущество векторных моделей, которое заключается в том, что им требуется на порядок меньшая память для хранения и меньшие затраты времени на обработку и представление.

При построении векторных моделей объекты создаются путем соединения точек прямыми линиями, дугами, полилиниями. Площадные объекты в векторных моделях задаются наборами линий.

Геоинформационные системы, работающие с векторными моделями, получили название векторных ГИС.

Конечно в реальности имеют дело не с абстрактными линиями и точками, а с объектами, содержащими линии и ареалы, занимающими пространственное положение, а также со сложными взаимосвязями между ними. Поэтому полная векторная модель данных отображает пространственные данные как совокупность следующих основных частей:

геометрические (метрические) объекты (точки, линии и полигоны);

атрибуты - признаки, связанные с объектами;

связи между объектами.

Векторные модели отображают все типы координатных моделей: точечные, линейные, площадные. Напомним, что типы координатных моделей определяются через базовый элемент *линия* следующим образом.

Точка определяется как линия нулевой длины. Линия - как линия конечной длины. Площадь представляется последовательностью связанных между собой линейных отрезков.

Векторные модели получают разными способами. Один из наиболее распространенных - векторизация сканированных (растровых) изображений. Она заключается в выделении векторных объектов со сканированного изображения и получении их в векторном формате. Другой способ - построение модели в режиме редактирования.



Некоторые объекты являются векторными по определению, например, дороги, границы соответствующего земельного участка, границы районов и т.д. Поэтому векторные модели обычно используют для сбора данных о границах и протяженных линейных или сетевых объектах.

В векторной форме упрощаются операции анализа и проектирования. Например, анализ сети, разработка маршрутов движения по сети дорог, анализ интенсивности движения и пр.

Векторные модели могут иметь любую заданную ширину линий, включая нулевую. В этом их преимущество перед растровыми, в которых точечный или линейный объект должен занимать группу ячеек. Это создает зависимость размеров растрового объекта от величины растра.

Поэтому точность векторных данных выше, чем растровых, так как векторные данные могут кодироваться с любой мыслимой степенью точности, которая ограничивается лишь возможностями метода внутреннего представления координат. Обычно для представления векторных данных используется 8 или 16 десятичных знаков (одинарная или двойная точность).

Только некоторые классы данных, получаемых в процессе измерений, соответствуют точности векторных данных. Это данные, полученные точной съемкой (координатная геометрия), карты небольших участков, составленные по топографическим координатам и политические границы, определенные точной съемкой.

Не все природные явления имеют характерные четкие границы, которые можно представить в виде математически определенных линий. Это обусловлено динамикой явлений или способами сбора пространственной информации. Почвы, типы растительности, склоны, место обитания диких животных - все эти объекты не имеют четких границ.

Поэтому для полного изучения явлений окружающей действительности нужно использовать как векторные так и растровые данные, используя их преимущества.

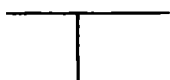
#### **10.4.2. Топологические модели и характеристики**

Большое количество пространственных объектов в геоинформатике имеют специфические связи. Наличие этих свойств определяет топологическое описание объектов. В таких случаях для описания объектов используют *топологические модели*.

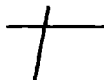


Рис. 10.2. Топологически родственные фигуры

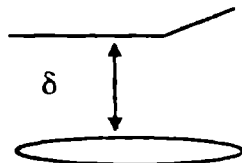
Трехвалентное



Четырехвалентное



а



б



в

Рис.10.3. Топологические характеристики пространственных объектов

**Топологические модели** - это модели, которые отражают взаимные связи между объектами, не зависящие от геометрических свойств объектов.

Напомним, что топологическими свойствами пространственных объектов называют такие, которые не изменяются при любых деформациях, производимых без разрывов или соединений. На рис. 10.2. представлены топологически родственные фигуры: прямоугольный четырехугольник, замкнутый контур произвольной формы, окружность, треугольник. Эти объекты (фигуры) имеют одинаковую топологию - одинаковые топологические свойства.

Рис. 10.3. а) показывает свойство *валентности узлов*. Рис.10.3 б) показывает характеристику *близость*. Рис. 10.3. в) показывает *связанность* графических объектов.

Другим примером топологически родственных фигур могут служить арифметические знаки сложения “+” и умножения “х”.

Топологическими характеристиками графические модели геоинформатики существенно отличаются от моделей САПР. Это различие просматривается в программно-технологическом обеспечении этих систем.

Топологические модели позволяют представлять элементы карты и всю карту в виде графов, т.е. применять теорию графов для описания пространственных объектов геоинформатики.

Теоретической основой топологических моделей служат алгебраическая топология и теория графов. В соответствии с алгебраической топологией координатные типы данных: площади, линии и точки называются 2-ячейками, 1-ячейками и 0-ячейками соответственно. Карта рассматривается как ориентированный двухмерный ячеечный комплекс.

Топологические характеристики вычисляются в ходе количественных преобразований моделей объектов, а затем с координатными данными хранятся в базе данных.

В геоинформатике топологическая модель определяется наличием следующих характеристик:

- **связанность векторов** - контуры, дороги и прочие векторы должны храниться не как независимые наборы точек, а как взаимосвязанные друг с другом объекты (рис. 10.3 в);
- **связанность и примыкание районов** - информация о взаимном расположении районов и об узлах пресечения районов (рис. 10.3 в);

- пересечение - информация о типах пересечений позволяет воспроизводить мосты и дорожные пересечения (рис. 10.3 а). Так Т - образное пересечение ( 3 линии) является трехвалентным, а Х - образное ( 4 линии сходятся в точке пересечения) называют четырехвалентным;
- близость - показатель пространственной близости линейных или ареальных объектов (рис. 10.3 б), оценивается числовым параметром, в данном случае символом  $\delta$ .

Топологические характеристики линейных объектов могут быть представлены визуально с помощью связанных графов. Граф сохраняет структуру модели со всеми узлами и пересечениями. При этом он напоминает карту с искаженным масштабом. Примером такого графа может служить схема метрополитена. Различие между картой метро и схемой метро отображает различие между картой и графом.

Узлы графа, описывающего картографическую модель, соответствуют пересечениям дорог, местам смыкания дорог с мостами, дорожным развязкам и т.п. Ребра такого графа описывают участки дорог и соединяющие их объекты. В данной модели длина ребер может не нести информативной нагрузки.

Топологические характеристики ареальных объектов представляют с помощью графов двух типов: покрытия и смежности.

**Граф покрытия** площадных объектов – это модель, которая упрощенно отображает реальное расположение объектов. Можно сказать, что граф покрытия соответствует видимой пространственной реальности.

Граф покрытия площадных объектов соответствует (топологически гомоморфен) контурной карте соответствующих районов. Ребра такого графа описывают границы между районами, а его узлы представляют точки смыкания районов. Степень вершины такого графа - это число районов, которые в ней смыкаются.

**Граф смежности** площадных объектов – это модель, которая отображает связи между пространственными объектами. Можно сказать, что граф смежности соответствует связям пространственной реальности. Эта модель искусственная, ее назначение - пространственный анализ связей между объектами.

Образно говоря, граф смежности это как бы вывернутый наизнанку граф покрытия. В нем районы отображаются узлами, а пара смыкающихся районов - ребрами. Используя такой граф можно

определить, является ли проходимой рассматриваемая территория, разделенная на проходимые или непроходимые участки.

Топологические характеристики описывают только одну сторону пространственных объектов – их связи. Поэтому для полноты описания они сопровождаются позиционной и описательной информацией.

Процесс наполнения информацией моделей объектов выглядит следующим образом.

На первом этапе осуществляют построение линейных и площадных объектов. После этого "создают" топологию. Эти процессы включают вычисление и кодирование связей между точками, линиями и ареалами.

Пересечения и связи имеют векторное представление. Топологические характеристики заносятся при кодировании данных в виде дополнительных атрибутов. Этот процесс осуществляется автоматически во многих ГИС в ходе дигитализации (картографических или фотограмметрических) данных.

Объекты связаны множеством отношений между собой. В свою очередь, отношения задают множества связей. Простейшие примеры таких связей: "ближайший к ...", "пересекает", "соединен с ...".

Каждому объекту можно присвоить признак, который представляет собой идентификатор ближайшего к нему объекта того же класса; так кодируются связи между парами объектов.

Таким образом, топологические характеристики являются важным свойством моделей реальных объектов в геоинформатике.

Все современные технологии сбора информации и геоинформационные технологии обработки данных должны включать этап ввода топологии в модели пространственных объектов.

### **10.5. Растровые модели**

Данный тип моделей уже упоминался в 10.4 при рассмотрении векторных моделей. Здесь мы даем его детальное рассмотрение.

Растровые модели это графические модели, описывающие пространственную реальность, в которых весь объект (исследуемая территория) отображается дискретно в площадные ячейки, образующие регулярную сеть [25, 32].

Регулярной сетью называют совокупность правильных геометрических фигур: квадраты, треугольники и др. При этом каждой

ячейке растровой модели соответствует одинаковый по размерам, но разный по характеристикам (цвет, плотность) участок (пиксель).

Каждая ячейка отображает реальный участок поверхности, который также характеризуется цветовыми характеристиками. Однако в ячейке растровой модели содержится одно значение, обобщающее цветовые и световые характеристики участка реальной поверхности. В теории обработки изображений эта процедура известна под названием пикселизации.

Если векторная модель дает информацию о том, *где* расположен тот или иной объект, то растровая - информацию о том, *что* расположено в той или иной точке территории. Это определяет основное назначение растровых моделей - непрерывное отображение поверхности.

В растровых моделях в качестве основной элементарной модели используют двухмерный элемент пространства - пиксель (ячейка). Упорядоченная совокупность атомарных моделей образует растр.

Растровые модели позволяют отображать полутона. Как правило, каждый элемент растра или каждая ячейка должны иметь лишь одно значение плотности или цвета. Это применимо не для всех случаев. Например, когда граница двух типов покрытий может проходить через центр элемента растра, элементу дается значение, характеризующее большую часть ячейки или ее центральную точку. Ряд систем позволяет иметь несколько значений для одного элемента растра.

### 10.5.1. Характеристики растровых моделей.

Для растровых моделей существует ряд характеристик: разрешение, ориентация, зоны, значение, положение.

**Разрешение** – это минимальный линейный размер наименьшего участка реальной поверхности, отображаемый одним пикселем.

Пиксели обычно представляют собой прямоугольники или квадраты, реже используются треугольники и шестиугольники. Более высоким разрешением обладает растр с меньшим размером ячеек. Высокое разрешение подразумевает обилие деталей, множество ячеек, минимальный размер ячеек.

**Значение** - элемент информации, хранящийся в элементе растра (пикселе). Поскольку при обработке применяют типизированные данные, то есть необходимость определить типы значений растровой модели.

Тип значений в ячейках растра определяется как реальным явлением, так и особенностями ГИС. В частности, в разных системах можно использовать разные классы значений: целые числа, действительные (десятичные) значения, буквенные значения.

Целые числа могут служить характеристиками оптической плотности или кодами, указывающими на позицию в прилагаемой таблице или легенде. Например, возможна следующая легенда, указывающая наименование класса почв: 0 - пустой класс, 1 - суглинистые, 2 - песчаные, 3 - щебнистые и т.п.

**Ориентацией** называют угол между направлением на север и положением колонок растра.

**Зона растровой модели** это соседствующие друг с другом ячейки, имеющие одинаковое значение и отображающие одно явление или объект. Зоной могут быть отдельные объекты, природные явления, ареалы типов почв, элементы гидрографии и т.п.

**Буферная зона** - вспомогательная зона, которая строится вокруг точечных, линейных или ареальных объектов на основе заданного правила или фиксированного значения.

Основные характеристики зоны - ее значение и положение.

**Положение** обычно задается упорядоченной парой координат (номер строки и номер столбца), которые однозначно определяют положение каждого элемента отображаемого пространства в растре.

При сравнении векторных и растровых моделей отмечено удобство векторных для организации и работы со взаимосвязями объектов. Тем не менее, используя простые приемы, например включая взаимосвязи в таблицы атрибутов, можно организовать взаимосвязи и в растровых системах..

Необходимо остановиться на вопросах точности отображения в растровых моделях. В растровых форматах в большинстве случаев неясно, относятся координаты к центральной точке пикселя или к одному из его углов. Поэтому точность привязки элемента растра определяют как  $1/2$  ширины и высоты ячейки

Растровые модели имеют следующие достоинства:

- растр не требует предварительного знакомства с явлениями, данные собираются с равномерно расположенной сети точек, что позволяет в дальнейшем на основе статистических методов обработки получать объективные характеристики исследуемых объектов. Благодаря этому растровые модели могут использоваться для изучения но-

вых явлений, о которых не накоплен материал. В силу простоты этот способ получил наибольшее распространение;

- растровые данные удобнее для обработки по параллельным алгоритмам и этим обеспечивают более высокое быстродействие по сравнению с векторными;
- некоторые задачи, например создание буферной зоны, гораздо проще решать в растровом виде;
- многие растровые модели позволяют вводить векторные данные, в то время как обратная процедура весьма затруднительна для векторных моделей;
- процессы растеризации много проще алгоритмически, чем процессы векторизации, которые зачастую требуют экспертных решений.

Наиболее часто растровые модели применяют при обработке аэрокосмических снимков для получения данных дистанционных исследований Земли.

**Метод группового кодирования.** Самый простой способ ввода растровых моделей - прямой ввод одной ячейки за другой. Недостатками данного подхода являются требования большого объема памяти и значительного времени для организации процедур ввода-вывода. Например, снимок искусственного спутника Земли (ИСЗ) Landsat имеет 74 000 000 элементов растра, что требует огромных ресурсов для хранения данных.

При растровом вводе информации в ГИС возникает проблема ее сжатия, так как наряду с полезной может попадать и избыточная (в том числе и бесполезная) информация. Для сжатия информации, полученной со снимка или карты, применяется кодирование участков развертки или метод группового кодирования, учитывающий, что довольно часто в нескольких ячейках значения повторяются.

Суть метода группового кодирования в том, что данные вводятся парой чисел, первое обозначает длину группы, второе - значение. Изображение просматривается построчно, и как только определен тип элемента или ячейки встречается впервые, он помечается признаком начала. Если за данной ячейкой следует цепочка ячеек того же типа, то их число подсчитывается, а последняя ячейка помечается признаком конца. В этом случае в памяти хранятся только позиции помеченных ячеек и значения соответствующих счетчиков.



Применение такого метода значительно упрощает хранение и воспроизведение изображений (карт), когда однородные участки (как правило) превосходят размеры одной ячейки.

Обычно ввод осуществляют слева направо, сверху вниз. Рассмотрим, например, бинарный массив матрицы (5 x 6) :

```

0 0 0 1 1 1
0 0 1 1 1 0
0 0 1 1 1 0
0 1 1 1 1 1
0 1 1 1 1 1.

```

При использовании метода группового кодирования он будет вводиться как :

3 0 3 1 2 0 3 1 3 0 3 1 2 0 5 1 1 0 5 1.

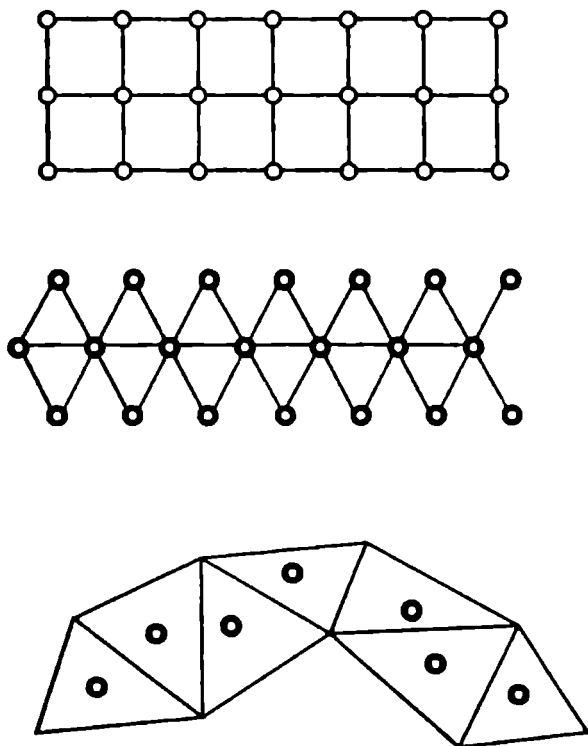
Вместо 30 необходимо только 20 элементов данных. В рассмотренном примере экономия составляет 30 %, однако на практике при работе с большими массивами бинарных данных она бывает гораздо больше.

Метод группового кодирования имеет ограничения и может использоваться далеко не во всех ГИС.

Элементы бинарной матрицы, т.е. растровой модели, могут принимать только два значения: "1" или "0". Эта матрица соответствует черно-белому изображению. На практике возможно полутоновое или цветное изображение. В этих случаях значения в ячейках растровой модели могут различаться по типам. Тип значений в ячейках растра определяется как исходными данными, так и особенностями программных средств ГИС.

В качестве значений растровых данных могут быть применены целые числа, действительные (десятичные) значения, буквенные значения.

В одних системах используются только целые числа, в других - различные типы данных. При этом ставится условие единства значений для отдельных растровых слоев. Целые числа часто служат кодами, указывающими на позицию в прилагаемой таблице или легенде.



**Рис.10.4. Регулярные и нерегулярные растровые модели**

Растровые модели делятся на регулярные мозаики, нерегулярные мозаики, (рис. 10.4).

Плоские регулярные мозаики бывают трех типов: квадрат (рис. 10.4) верхняя модель, треугольник и шестиугольник (рис. 10.4) средняя модель. Квадрат - самая удобная модель, так как позволяет относительно просто проводить обработку больших массивов данных. Треугольные мозаики служат хорошей основой для создания выпуклых (сферических) покрытий.

Среди нерегулярных мозаик чаще всего используют треугольные сети неправильной формы (Triangulated Irregular Network - TIN) и полигоны Тиссена (рис. 10.4). Сети TIN удобны для создания цифровых моделей отметок местности по заданному набору точек. Они применяются как в растровых, так и в векторных моделях.

Следует отметить, что полигоны Тиссена являются продолжением и развитием работ российских ученых Вороного и Делоне, о которых речь пойдет ниже.

## 11. Геоинформационные системы (ГИС)

### 11.1. Развитие и определение ГИС

В настоящее время существует множество определений ГИС. Эта особенность обусловлена тем, что исторически ГИС развивались длительной время. Естественно, что они совершенствовались и, по мере их изменения, менялись их основные функции и свойства. Все это приводило к появлению новых определений ГИС, учитывающих их новые свойства и новые особенности [4, 8, 11, 20, 44-46, 51, 56, 74, 75, 82, 98, 105].

Первые ГИС назывались "Географические информационные системы". Их определяли как комплексный блок компьютеров, предназначенных для ввода, хранения, анализа и вывода пространственно связанных данных. Эти ГИС ведут свое происхождение от Географической информационной системы Канады, построенной на базе первых больших ЭВМ и пакетной системы обработки данных. Эта первая ГИС была разработана в начале 1960-х гг. задолго до появления недорогих персональных компьютеров и доступных для пользования баз данных. В это время геоинформатика как наука еще не сформировалась.

Несмотря на технические ограничения таких систем, было выявлено, что определенные виды анализа карт и инвентаризации, могут выполняться на ЭВМ значительно эффективней, чем вручную.

Следующим этапом развития ГИС считают начало 80-х годов. Этот этап связывают с появлением реляционной модели данных, разработанной Коддом еще в 1969-70 гг. и разработанных на ее основе реляционных баз, заменивших иерархические базы данных.

Увлечение базами данных привело к тому, что ГИС стали называть "компьютеризированной базой данных" для хранения географической и тематической информации.

Однако базы данных предназначаются в основном для хранения, а не для обработки информации. Поэтому ГИС продолжали развиваться.

После разработки и использования специального программного обеспечения для решения задач геоинформатики появились новые определения ГИС, как "систем программного обеспечения" (без включения в них технологических возможностей). В этот период геоинформатика как наука завершала свое формирование.

Увлечение автоматизацией кадастра, в первую очередь городского, привело к появлению названия - городские информационные системы.

Следующим этапом развития ГИС следует считать использование программного обеспечения систем автоматизированного проектирования (САПР) для работы с графической информацией и построения карт. Это середина 80-х годов. В это время еще не были созданы специализированные программные продукты для задач ГИС и работы по автоматизированному составлению карт велись с помощью программных пакетов САПР. Следует упомянуть популярный продукт того времени "Автокад" фирмы Autodesk.

Главным достижением программного обеспечения САПР явилось создание системы послойного представления графической информации (чертежей, карт). Графическая информация типизировалась (см. гл.2) и типизированные данные помещались в отдельный слой. Каждый слой напоминал бы прозрачную бумагу. Совокупность слоев создавала привычную картину графической информации, но в отдельности каждый слой было удобно обрабатывать независимо от других.

Другим достижением САПР было введение так называемых "блоков" для получения проектных документов. Блоками называли типовые повторяющиеся элементы, хранящиеся в базе данных. Блок создавался независимо и вставлялся в заданные точки чертежа неограниченное число раз. Это существенно сокращало объем данного чертежа при его хранении в базе данных.

В случае редактирования блока, редакция автоматически происходила во всех частях чертежа, в которых этот блок был вставлен. Этим существенно сокращались редакционные работы графического документа. Позже механизм блоков послужит основой создания библиотек условных картографических знаков.

В начале 90-х годов стали появляться интегрированные программные продукты и интегрированные информационные системы. Информатика становится основой подготовки различных специалистов в области обработки информации [6, 25-28, 55, 59-60, 83-86, 92, 114].

Кроме того, появилась потребность в информационных системах, позволяющих осуществлять глобальную интеграцию различных видов информации. Появились новые информационные технологии и системы.

Все это привело к необходимости создания ГИС как автоматизированной интегрированной информационной системы [8, 98].

На рис. 11.1. приведена структура ГИС как автоматизированной интегрированной информационной системы (АИИС). Она включает следующие подсистемы:

- подсистему сбора,
- подсистему хранения данных (чаще всего это база данных или экспертная система),
- подсистему обработки данных (моделирования),
- подсистему представления информации,
- телекоммуникационную подсистему.

Наличие этих подсистем определяет различные аспекты интеграции данных и методов обработки.

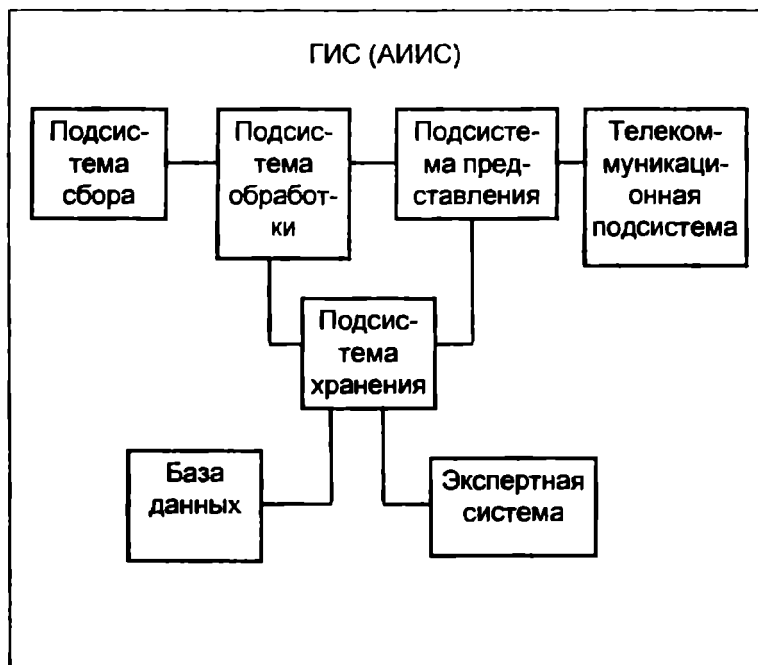


Рис.11.1. Структура ГИС как автоматизированной интегрированной информационной системы

Перечислим основные аспекты интеграции: интеграция исходных данных, интеграция технологий сбора, интеграция данных для хранения и моделирования, интеграция технологий обработки, интеграция технологий хранения, интеграция данных для представления и передачи, интеграция технологий представления информации.

В современных ГИС следует различать два дополняющих друг друга вида интеграции:

интеграцию технологий;

интеграцию данных.

Интеграция технологий в ГИС заключается в разработке комплекса взаимосвязанных технологий на основе некоей базовой технологии.

Интеграция данных означает в ГИС, что для создания информационной основы данной информационной системы выбирают определенный класс данных, а все остальные типы данных преобразуются применительно к свойствам этого класса.

Интеграция технологий в ГИС осуществляется на основе технологий САПР, а интеграция данных в ГИС осуществляется на основе географической информации.

Таким образом, возникнув как специализированная или прикладная географическая информационная система, ГИС в процессе эволюции стала обобщением автоматизированных информационных систем и перешла в категорию интегрированных предметных информационных систем.

Напомним, что предметными информационными системами (базами данных) называют системы, относящиеся к предметной области. Прикладными информационными системами (базами данных) называют более узкий класс систем, решающих прикладные задачи.

Объектом исследования ГИС стала не только география или географическая информация, а все процессы и явления, происходящие на земной поверхности. Именно поэтому современная ГИС является геоинформационной системой. Отсюда следует, что понятия «геоинформационная система» и «географическая информационная система» - неравноценны. Вторые являются частным случаем первых.

Рассмотренный процесс эволюции ГИС дает возможность сформулировать следующее определение для нее.

*ГИС - интегрированная автоматизированная информационная система, предназначенная для обработки пространственно-временных данных, основой интеграции данных, в которых является*

*географическая информация, а основой интеграции технологий обработки в ГИС являются технологии САПР.*

## 11.2. Особенности ГИС

Для дальнейшего описания ГС рассмотрим ряд ее особенностей.

1. Основой интеграции данных в ГИС является географическая информация, но большинство задач решаемых в ГИС далеки от географических.

2. Основой интеграции технологий в ГИС являются технологии САПР, но решаемые задачи значительно шире проектных.

3. По формальному определению ГИС относится к информационным системам (т.е. к системам АСИС хранения информации), но по своему функциональному назначению ГИС относятся к классу систем обработки данных и управления (СОДУ).

4. Современная ГИС рассчитана не просто на обработку данных, а на проведение во многих ситуациях экспертных оценок. Другими словами ГИС должна включать в свой состав экспертную систему или экспертную технологию.

Данные, которые обрабатывает и хранит ГИС, имеют не только пространственную, но и временную характеристику. По временной характеристики информация, хранимая в ГИС, подразделяется на:

- долгосрочную (десятки лет хранения);
- среднесрочную (годы);
- годовую и сезонную;
- оперативную.

В ГИС осуществляется комплексная обработка информации от сбора данных до ее хранения, обновления и представления, поэтому следует рассмотреть ГИС с различных позиций.

Как системы управления ГИС -предназначены для обеспечение процесса принятия решений по оптимальному управлению землями и ресурсами, городским хозяйством, организации транспорта и розничной торговли, использованию океанов или других пространственных объектов [68].

Отличие от автоматизированных систем управления АСУ в ГИС появляется множество новых технологий пространственного анализа данных. В силу этого ГИС является мощным средством преобразования и синтеза разнообразных данных для задач управления, в котором присутствуют эффективные методы для точного и быстрого применения картографической информации.



Опишем применение ГИС в различных областях.

Как автоматизированные информационные системы ГИС объединяют ряд технологий или технологических процессов известных информационных систем типа АСНИ, САПР, АСИС.

Как "geo" системы ГИС включают технологии (прежде всего технологии сбора информации) таких систем как: Географические информационные системы (ГИС), Системы картографической информации (СКИ), Автоматизированные системы картографирования (АСК), Автоматизированные фотограмметрические системы (АФС), Земельные информационные системы (ЗИС), Автоматизированные кадастровые системы (АКС) и т.п.

Как системы, использующие базы данных, ГИС характерны широким набором данных, собираемых с помощью разных методов и технологий. При этом следует подчеркнуть, что они объединяют в себе как базы данных обычной (цифровой) информации, так и графические базы данных.

Как системы моделирования ГИС используют максимальное количество методов и процессов моделирования, применяемых в других автоматизированных системах.

Как системы получения проектных решений ГИС во многом используют методы автоматизированного проектирования и решают ряд специальных проектных задач, которые в типовом автоматизированном проектировании не встречаются.

Как системы представления информации ГИС являются развитием автоматизированных систем документационного обеспечения с использованием современных технологий MultiMedia. Это определяет значительно большую наглядность по сравнению с обычными картами. Технологии представления данных позволяют оперативно получать визуальное представление картографической информации с различными нагрузками, переходить от одного масштаба к другому, получать атрибутивные данные в табличной или графовой формах.

Как интегрированные системы ГИС являют собой пример современной интегрированной системы, созданной на основе интеграции технологий на основе технологий САПР и интеграции данных на основе географической информации, объединения различных методов и технологий в единый комплекс.

По территориальному признаку, т.е. по сфере действия ГИС на тот или иной регион вводят понятие территориального уровня использования ГИС

В России выделяют следующие территориальные уровни применения ГИС.

Глобальный уровень - Россия на глобальном и евразийском фоне масштаб 1 : 45 000 000 - 1 : 100 000 000.

Всероссийский уровень - вся территория страны, включая прибрежные акватории и приграничные районы, масштаб 1 : 2 500 000 - 1 : 20 000 000.

Региональный уровень - крупные и природные экономические регионы, субъекты федерации М 1 : 500 000 - 1 : 4 000 000.

Локальный уровень - области, районы, национальные парки, ареал кризисных ситуаций - 1 : 50 000 - 1 000 000.

Муниципальный уровень - города, городские районы, пригородные зоны. М 1 : 50 000 и крупнее.

Как прикладные системы ГИС не имеют себе равных по широте, т.к. применяются в транспорте, навигации, геологии, географии, военном деле, топографии, экономике, экологии и т.д. Как следствие широких возможностей применения ГИС для решения различных специальных задач на их основе интенсивно развивается тематическое картографирование. Именно эффективность последнего обеспечивает разнообразное решение задач в разных отраслях при использовании интеграции данных на основе картографической информации.

Можно считать, что современные ГИС интегрировали в себя технологии практически всех других автоматизированных систем топографо-геодезического производства. Разумеется это не исключает существование "чисто" Географических информационных систем - ГИС, однако в более общем случае будем понимать под термином ГИС геоинформационные системы.

При этом следует подчеркнуть, что ГИС как системы массового пользования позволяют использовать картографическую информацию на уровне деловой графики, что делает их доступными не только специалисту географу, а любому школьнику или бизнесмену.

Именно поэтому применение ГИС -технологий не сводится к созданию карт, а лишь использует картографические данные для принятия решений во многих областях.

Основой создания ГИС служат так называемые инструментальные программно-технологические комплексы. В отличие от систем "под ключ" инструментальные допускают настройку под конкретные запросы пользователя, что может приводить к различию в организа-

ции структур данных, проектированию баз данных, функциональному назначению даже на базе родственных систем.

Инструментальные системы позволяют отображать различные данные, имеющие пространственную привязку и создавать интегрированные геоинформационные технологии, прикладные геоинформационные системы, цифровые картографические системы.

Основу процессов обработки информации в ГИС составляет цифровое моделирование. Использование такого типа моделирования позволяет осуществлять векторно-топологическое моделирование, буферизацию объектов, анализ сетей, построение ЦММ и т.д.

Основным назначением ГИС следует считать формирование знаний о процессах и явлениях на земной поверхности и применение этих знаний для решения практических задач во всех сферах человеческой деятельности.

Подводя итог, следует констатировать, что ГИС в настоящее время представляют собой современный тип интегрированной информационной системы, применяемой в разных направлениях. Она отвечает требованиям глобальной информатизацией общества. ГИС является системой способствующей решению управленческих и экономических задач на основе средств и методов информатизации, т.е. способствующей процессу информатизации общества в интересах прогресса.

ГИС как система и ее методология совершенствуются и развиваются, ее развитие осуществляется в следующих направлениях:

- развитие теории и практики информационных систем;
- изучение и обобщение опыта работы с пространственными данными;
- исследование и разработка концепций создания системы пространственно-временных моделей;
- совершенствование технологии автоматизированного изготовления электронных и цифровых карт;
- разработки технологий визуальной обработки данных;
- разработки методов поддержки принятия решений на основе интегрированной пространственной информации;
- интеллектуализации ГИС.

К недостаткам ГИС как информационной системы следует отнести слабые программно-технологические реализации по обработке данных и принятию решений при потенциально больших возможностях.

## 12. Цифровое моделирование и цифровые модели

### 12.1. Основные понятия

Развитие автоматизированных методов обработки пространственной информации привело к появлению нового направления в моделировании - цифрового моделирования.

Термин «цифровое моделирование» - синоним термина «компьютерное моделирование». В геоинформатике оно заключается в использовании возможностей математических методов и программных средств для моделирования объектов земной поверхности.

Основным элементом данного моделирования является цифровая модель (местности) ЦМ(М), которая может быть получена с помощью разнообразных технологий.

В отличие от всех рассмотренных выше теоретических моделей цифровые являются их формой представления для обработки на ЭВМ. Из этого следует, что цифровая модель может иметь в качестве структурной основы иерархическую, реляционную, сетевую или комплексную модель.

Цифровые модели хранятся двумя способами в базах данных или независимо в виде файловых структур. Наибольшее распространение цифровые модели нашли в геоинформатике, строительстве, архитектуре.

В настоящее время существует множество определений ЦММ. Это объясняется их эволюцией и узкоспециальными областями применения. Тем не менее, можно выделить общие признаки ЦММ с помощью методов абстракции для построения структур моделей.

Структура ЦММ, построенная на основе обобщения, (с использованием принципа "цифровая модель есть часть ..." [98, 124]) отражает четыре основных свойства модели, вытекающие из ее определения.

1. Как *модель объекта* конкретной предметной области она должна быть специализированной моделью, т.е. содержать специальную информацию о данной предметной области. Из этого следует необходимость определена логической структуры модели на основе информационных единиц данной области.

2. Как *модель данных* ЦММ определяется на известном классе моделей, т.е. имеет вполне определенную структуру и содержит в своей основе одну или несколько теоретических моделей данных.

Это означает, что ЦММ должна быть организована в соответствии с правилами построения и использования класса моделей.

3. ЦММ как *модель цифровая* должна быть оптимально организована и удобна при работе на ЭВМ. Это означает, что для полной реализации модели должна быть определена ее физическая структура.

4. Как *элемент базы данных* ЦММ предназначена для моделирования, многократного использования, анализа и решения различных задач. Отсюда следует, что ЦММ должна содержать дополнительную информацию для ее многократного использования.

Из четвертого свойства вытекает, что ЦММ должна быть информативно переопределена по отношению к одиночной модели объекта, Это означает, что она должна содержать свойства подкласса (группы) объектов, а не только одного объекта. Это свойство необходимо учитывать при организации технологии систем, использующих цифровое моделирование

Второе и третье свойство определяет синтаксис модели, первое семантику, четвертое прагматику. Таким образом, цифровая модель содержит все три семиотические части и является полной моделью.

Особенность цифровой модели в ее ресурсности [59, 104]. Эффект от такой модели возрастает с течением времени, по мере заполнения базы данных и ее оптимального использования. В ресурсной модели информация накапливается не только в виде суммы данных о *единичных объектах*, но и в виде *взаимосвязанных методов преобразования информации* и *накапливаемой суммы свойств и связей классов объектов*.

Обобщенное описание цифровых моделей местности должно выполняться на уровне *типов*, т.е. для этого необходимы предварительный анализ и последующая максимальная *титизация* пространственных объектов.

Индивидуальные свойства конкретного объекта должны выражаться на уровне *знаков*. Такое сочетание индивидуального и обобщенного описания в теории моделей данных известно под названием *классификация*. Следовательно, для полного создания ЦММ необходимо разработать систему классификации или некий классификатор.

## 12.2. Математическая основа цифровых моделей

Одной из задач геоинформатики и цифрового моделирования является определение местоположения объектов на поверхности Земли.

Традиционно для решения подобных задач использовались карты или геодезические методы. Особенность использования картографической информации в том, что преобразование пространственных (трехмерных) данных с поверхности Земли в плоские (двухмерные) координаты карты всегда сопряжено с искажениями. Картографическая информация была основной входной информацией в ГИС первого поколения и по настоящее время остается основной формой представления цифровых моделей ГИС. Поэтому очень важно исследовать особенности применения картографической информации при цифровом моделировании.

Карта может быть рассмотрена как двухмерная модель, отображающая трехмерную поверхность Земли. Построение карты связано с моделированием (отображением) трехмерной поверхности в плоскость.

Следует учесть, что в геодезии принято измерения, выполненные на физической поверхности, переносить на математическую поверхность. Эту математическую поверхность называют математической основой. Ее выбирают как модель, наиболее близкую к физической, которая может быть описана соответствующими уравнениями. В настоящее время в качестве математической основы используют общеземной эллипсоид и референц-эллипсоиды.

Особенность моделирования точек земной поверхности в том, что Земля представляет собой тело неправильной геометрической формы. Следовательно, прежде чем отображать точки земной поверхности на карту необходимо выбрать приемлемую модель земной поверхности на которую можно отнести реальные точки. Другим словами, невозможно создать модель карты без математической модели поверхности Земли.

Исследованием вопросов отображения точек земной поверхности на топографические карты занимается специальная наука математическая картография [3, 4].

Представления о фигуре Земли менялись по мере ее изучения и развития науки. С VI века до н. э. до конца XVII века Землю принимали за шар. Такое тело называли сферондом.

С конца XVII века до второй половины XIX века в качестве формы Земли использовали фигуру, близкую к эллипсоиду вращения. Со второй половины XIX века до сороковых годов XX столетия в качестве модели Земли используют трехосный эллипсоид. Трехосный эллипсоид является лишь приближенным отображением более сложной формы геоида. Геоид определяется с помощью понятия урвенной поверхности.

**Урвенной поверхностью** называют поверхность, каждая точка которой ортогональна к вектору силы тяжести данного небесного тела.

Набор урвенных поверхностей задает систему высот. В каждой стране начальная урвенная поверхность имеет свой начальный по высоте пункт, поэтому системы высот в разных странах различаются. В России, принята Балтийская система высот.

**Балтийская система высот** это система, для которой начальная урвенная поверхность проходит через нуль-пункт Кронштатского футштока.

**Поверхностью геоида** называют урвенную поверхность, проходящую через точку начала отсчета высот.

Строгое определение геоида связано со знанием строения земной коры. Следует отметить, что на Земле имеются так называемые гравитационные аномалии, вследствие чего урвенная поверхность не может быть удобной математической моделью. Поэтому на практике используют понятие квазигеоида. Эту фигуру можно определить без гипотез о строении земной коры, она имеет более гладкую (регулярную) поверхности по сравнению с геоидом.

**Поверхность квазигеоида** совпадает с поверхностью геоида на морях и океанах и отступает от неё до 2 метров на континентальной части Земли.

Здесь следует напомнить, что центры масс суши и мирового океана не совпадают.

Поверхность Земли имеет сложную форму. При общей площади ее поверхности 510 млн. км<sup>2</sup> 71 % приходится на дно морей и океанов и 29 % - на сушу. Это дает основание считать, что земная поверхность состоит из двух резко отличающихся морфологических элементов - материков и океанов.

С помощью методов дистанционного зондирования удалось установить, что Земля имеет грушевидную форму.

**Общеземным эллипсоидом** называют эллипсоид вращения, плоскость экватора и центр которого совпадает с плоскостью эквато-

ра и центром масс Земли и наилучшим образом аппроксимирует поверхность геоида (квазигеоида).

Референц-эллипсоидом называют эллипсоид вращения, на поверхность которого отображаются материалы астрономо-геодезических работ, топографических съемок и геодезические.

В СССР до 1946 г. использовался эллипсоид, полученный Ф.Бесселем. Его параметры:

большая полуось  $a$  - 6377397.155 м.

малая полуось  $b$  - 6356079 м.

сжатие  $(a-b)/f$  - 1:299.1528128.

В 1946 г. для обязательного использования был введен эллипсоид, вычисленный в ЦНИИГАиК в 1940 г. под руководством Ф.Н. Красовского при участии А.А. Изотова. Его параметры:

большая полуось  $a$  - 6378245 м.

малая полуось  $b$  - 6356863.0188 м.

сжатие  $(a-b)/f$  - 1:298.3

За рубежом, особенно в программном обеспечении зарубежных ГИС, широко используется референц - эллипсоид Кларка [7, 9]. Существует два вида Кларк I и Кларк II. Их параметры приведены в таблице:

Таблица. Параметры эллипсоида Кларка

	Кларка I	Кларка II
Большая полуось $a$	6378206 м	6378249 м
Малая полуось $b$	6356584 м	6356515 м
Сжатие $(a-b)/f$	1:294.9786982	1:293.465

### 12.3. Геодезическая система координат

Как известно географические координаты (широта, долгота) это всего лишь двумерные характеристики пространственных объектов. Они не могут эффективно использоваться для точных расчетов и при решении инженерных задач. Более точное положение дают геодезические координаты которые задают трехмерные координаты объекта.

Кроме математической основы при определении точек земной поверхности используют геодезическую основу. Она используется для того, чтобы привязать математическую модель земной поверхности к конкретной реальной поверхности с целью обеспечения наибольшей точности измерений именно в этой части поверхности.

Элементами геодезической основы являются опорные пункты, определённые в системе геодезических координат, принятой в дан-



ном государстве, и координатные сетки, связанные с этими опорными пунктами.

Геодезические системы координат включают:

- параметры референц-эллипсоида (величина большой полуоси  $a$  или малой  $b$ , сжатие  $a$  или эксцентриситет  $e$ );
- высоту геоида над референц-эллипсоидом в начальном пункте;
- исходные геодезические даты (геодезические широта и долгота начального пункта, азимут на ориентирный пункт).

В работах по геодезии, топографии и картографии, выполняемых в России, используется эллипсоид Красовского, начальный пункт Пулково; превышение геоида над референц-эллипсоидом в начальном пункте равно нулю.

При создании карт на российские дальневосточные регионы иногда применяется система высот Охотского моря. В процессе выполнения картосоставительских работ определяют геодезическую систему координат и систему высот, которые были приняты при создании исходного картографического материала. Это выявляется по формулярам листов карт или по литературно-описательным источникам.

В настоящее время разработаны специальные программы, позволяющие осуществить преобразование одной геодезической системы координат в другую.

## 12.4. Системы координат

Координатные данные, составляющие один из основных классов геоинформационных данных, используют для указания местоположения на земной поверхности.

Для отображения положения точек поверхности на плоскости используют различные математические модели поверхности и различные системы координат. На практике применяют два основных типа координат: плоские и сферические. Реже применяют криволинейные или полярные.

Выбор системы координат зависит от размеров исследуемых участков поверхности, как следствие, от влияния кривизны Земли. При изображении небольших участков Земли часть уровенной поверхности можно принять за плоскость. Такими участками будут участки до 20 км длиной и площадью до 400 км<sup>2</sup>.

В этих случаях применимы плоские координаты. Плоские декартовы координаты определяются заданием двух осей. При этом

обычно координата  $X$  указывает на восток,  $Y$  - на север. Задают масштабные отрезки. Упорядоченная пара  $(X, Y)$  определит положение точки в заданной системе.

**Плоские полярные координаты** используют расстояние от начала координат ( $r$ ) и угол ( $\varphi$ ) от фиксированного направления. Направление обычно фиксируется на север, а угол отсчитывается по часовой стрелке от него. Полярные координаты удобны при проведении измерений от какой-либо заданной точки, например когда используются данные таких источников, как радарные съемки.

При необходимости учета кривизны Земли применяют пространственные системы координат.

Для определения географической системы координат (разновидности сферической системы) введем следующие понятия:

**плоскость земного экватора** - проходит через центр Земли перпендикулярно к оси вращения;

**плоскость географического (астрономического) меридиана** - проходит через ось вращения Земли и отвесную линию в точке земной поверхности;

**меридиан** - линия пересечения плоскостей географических меридианов с земной поверхностью;

**параллель** - линия, образованная пересечением плоскости, параллельной плоскости земного экватора, с поверхностью Земли.

Положение точки в географической системе координат определяется широтой ( $\varphi$ ) и долготой ( $\lambda$ ).

**Широта** - это угол между точкой и экватором вдоль меридиана<sup>1</sup>. Она изменяется от  $-90^\circ$  (южный полюс) до  $+90^\circ$  (северный полюс).

**Долгота** - это угол в плоскости экватора между меридианом точки и главным (нулевым) меридианом, проходящим через Гринвич (Англия). Она изменяется от  $-180^\circ$  (западная долгота) до  $+180^\circ$  (восточная долгота).

Основными понятиями этой системы координат являются:

**меридиан** - линия постоянной долготы;

**параллель** - линия постоянной широты;

**большой круг** - воображаемый круг на земной поверхности, образованный плоскостью, проходящей через центр земного шара;

---

<sup>1</sup> Такое определение широты не годится для эллипсоида.

**малый круг** - воображаемый круг на земной поверхности, образованный плоскостью, не проходящей через центр земного шара.

Рассмотренные системы координат носят в большей степени теоретический характер. На практике используют более широкий набор систем координат: геоцентрические, топоцентрические, полярные геодезические, эллиптические и др.

## 12.5. Картографические проекции

При составлении карт пространственное (трехмерное) положение точек отображается в плоском (двухмерном) представлении. Для отображения положения точек трехмерной поверхности на плоскости применяют различные математические методы.

Картографические проекции выполняют функции представления на плоской поверхности объектов, расположенных на поверхности Земли. Более точно они отображают не сами объекты, а их образы, перенесенные на эллипсоид или геоид, т.е. некую регулярную поверхность, поскольку реальная поверхность и форма Земли не является правильной геометрической фигурой.

Земной эллипсоид задается набором параллелей и меридианов, которые создают некую основу, на которую переносятся объекты земной поверхности, имеющие к тому же разные высотные отметки.

Существуют различные виды проекций и соответственно проекционных преобразований (около 1000), но все они основаны на формировании специфического рисунка параллелей широты и меридианов долготы эллипсоида на выравшиваемую или разворачиваемую поверхность. Картографические проекции образуют исходя из разных принципов. Одним из них является принцип геометрического построения.

В геометрии тремя наиболее простыми разворачиваемыми поверхностями являются плоскость, цилиндр, конус.

Это определило три семейства картографических проекций: азимутальные, цилиндрические, конические.

Следует четко представлять, что любой способ отображения пространственной криволинейной поверхности в плоскую влечет ошибки. Поэтому неизбежно от типа и класса любой применяемой картографической проекции неизбежны ошибки или искажения, привносимые этой проекцией.

В сущности, любая карта это картографическая проекция с нанесенной на нее картографической информацией. В идеальном случае карта, как плоский образ объекта, имеет четыре характеристики,

определяющие ее соответствие или истинность исходному пространственному объекту:

конформность или равноугольность;  
 эквивалентность или равновеликость;  
 равнодистантность или равнопромежуточность;  
 истинность направлений.

Никакая картографическая проекция не может быть истинной по всем перечисленным четырем характеристикам. Поэтому проекция планируется быть истинной только по ограниченным характеристикам, пренебрегая истинностью других, или компромиссной.

Карта как образ пространственных объектов в первую очередь отображает их топологию, во вторую их метрические характеристики.

**Конформность (равноугольность)** характеризует истинность формы, когда проекция сохраняет форму любой географической области маленького размера.

**Эквивалентность (равновеликость)** – характеристика равенства площадей, что означает равенство *масштабов площадей* частей карты. Созрание эквивалентности приводит к неточному трансформированию углов между линиями в заданной точке.

**Равнодистантность (равнопромежуточность, эквидистантность)** – характеристика истинного измерения расстояний. Она требует неизменности *масштабов расстояний*. Невозможно обеспечить равнопромежуточность для всех линий, поэтому на карте выбирают эталонные линии (экватор, меридиан), для которых эта характеристика истинна.

**Истинность направлений** характеризуется вектором между двумя точками, который пересекает эталонные линии (экватор, меридиан) под постоянным углом (азимутом). Азимут – угол, измеряемый от меридиана с севера на восток по часовой стрелке. Линия постоянного или одинакового направления называется румбовой линией. Характеристика направления используется при проложении курсов в задачах навигации.

В настоящее время существует множество подходов построения картографических проекций. Это определяет разнообразие способов их классификации. Картографические проекции классифицируют по разным признакам [7]:

- по характеру искажений;
- по способам получения проекций

- по ориентировке картографической сетки в зависимости от положения точки полюса принятой системы координат;
- по виду нормальной картографической сетки линий;
- по виду общих уравнений картографических проекций;
- и другим.

Классификации проекций по множеству признаков подробно рассмотрена в теории математической картографии [7] Мы рассмотрим два первых подхода.

По характеру искажений картографические проекции могут быть равноугольными, равновеликими и произвольными (в частных случаях равнопромежуточными).

Равноугольными проекциями называют такие, в которых отсутствуют искажения углов и азимутов линейных элементов. Эти проекции (conformal projection) сохраняют без искажений углы и формы малых объектов, но в них резко деформируются длины и площади объектов. В математике такие преобразования называют конформными.

Этот вид проекций приемлем для прокладки маршрутов транспортных средств

Равновеликими проекциями называются такие, в которых площади соответствующих областей на поверхности эллипсоида (сферы) и на плоскости тождественно равны (пропорциональны). Эти проекции (equivalente projection) не искажают площадей, но в них искажены углы и формы объектов.

Этот вид проекций приемлем для определения площадей и землепользования.

Произвольные проекции (arbitrary projection) имеют искажения углов, площадей и длин, но эти искажения распределены по карте например так, что минимальные искажения имеются в центральной части и возрастают к краям. Среди произвольных проекций выделяют *равнопромежуточные* (equidistant projection), в которых искажения длин отсутствуют по одному из направлений: вдоль меридиана или вдоль параллели.

Равнопромежуточными называются проекции, сохраняющие длины по одному из главных направлений. Наиболее часто к ним относят проекции с ортогональной картографической сеткой. В этих случаях главными будут направления вдоль меридианов и параллелей. Соответственно определяются равнопромежуточные проекции вдоль одного из этих направлений

Картографические проекции задаются с помощью систем аналитических выражений. На практике эти аналитические выражения необходимо запрограммировать, упорядочить и поместить в специальный программный блок.

Этот программный блок математических процедур, осуществляющих переход от одной картографической проекции к другой или от пространственной системы к картографической проекции, носит название проекционных преобразований.

Число проекционных преобразований в блоках современных геоинформационных систем различно: в системе ER Mapper их свыше 700, в ГеоГраф - около трех десятков.

Преимуществом ГИС как интегрированной системы является возможность выполнения проекционных преобразований не только при работе с картографической информацией но и с любой другой. Например, возможно трансформирование космического снимка с дополнительным преобразованием его в заданную картографическую проекцию.

Такая технология позволяет оперативно создавать фотокарты или комбинированные карты, в которых на картографическую основу наносятся оперативные данные космических съемок. Эта возможность предоставляется в пакетах ГИС, в первую очередь связанных с обработкой данных дистанционного зондирования.

По способу построения выделяют три больших класса проекций, включающие разновидности: конические, цилиндрические, азимутальные.

**Конические проекции** (conical projection) образуются на основе проектирования земной поверхности на боковую поверхность конуса, определенным образом ориентированного относительно земного шара (эллипсоида). В прямых конических проекциях оси земного шара и конуса совпадают. При этом конус берется или касательный, или секущий.

После проектирования боковая поверхность конуса разрезается по одной из образующих и разворачивается в плоскость. При проектировании по методу линейной перспективы получают перспективные конические проекции, обладающие только промежуточными свойствами по характеру искажений.

Другой метод образования конических проекций - аналитический. В его основу положены уравнения проекций, вытекающие из их определения и формулы общей теории искажений. В конических проекциях имеются две постоянные проекции "α" и "с". Постоянная

" $\alpha$ " равняется синусу широты стандартной параллели или, что то же самое, синусу угла при вершине конуса.

В зависимости от размеров изображаемой территории в конических проекциях принимаются одна или две параллели, вдоль которых сохраняются длины без искажений. Одна параллель (касательная) принимается при небольшом протяжении по широте; две параллели (секущие) - при большом протяжении для уменьшения уклонов масштабов от единицы. В литературе их называют стандартными параллелями.

Коническая проекция данной группы вполне определяется, если заданы постоянные проекции или любые величины, с ними связанные. Это могут быть широты стандартных или крайних параллелей. В последнем случае может быть введено дополнительное условие, чтобы масштабы на крайних параллелях и на параллели с наименьшим масштабом были равны по абсолютной величине.

Особенностью всех конических проекций является то, что их центральные линии совпадают со средними параллелями. Следовательно, конические проекции удобны для изображения территорий, расположенных в средних широтах или в южных широтах, асимметрично к экватору, и значительно вытянутых по долготе. Именно поэтому многие карты Советского Союза составлены в этих проекциях.

Азимутальные проекции (azimuthal projection) образуются на основе проектирования земной поверхности на плоскость, определенным образом ориентированного относительно земного шара (эллипсоида). В них параллели (альмукуантараты) изображаются концентрическими окружностями, а меридианы (вертикалы) - пучком прямых, исходящих из центра

Углы между меридианами проекции равны соответствующим разностям долгот. Промежутки между параллелями определяются принятым характером изображения (равноугольным или другим) или способом проектирования точек земной поверхности на картинную плоскость.

Нормальная сетка азимутальных проекций ортогональна. Их можно рассматривать как частный случай конических проекций, в которых  $\alpha=1$ .

Применяются прямые, косые и поперечные азимутальные проекции, что определяется широтой центральной точки проекции, выбор которой зависит от расположения территории. Меридианы и параллели в косых и поперечных проекциях изображаются кривыми

линиями, за исключением среднего меридиана, на котором находится центральная точка проекции. В поперечных проекциях прямой изображается также экватор: он является второй осью симметрии.

В зависимости от искажений азимутальные проекции подразделяются на равноугольные, равновеликие и с промежуточными свойствами. В проекции масштаб длин может сохраняться в точке или вдоль одной из параллелей (вдоль альмукантарата). В первом случае предполагается касательная картинная плоскость, во втором - секущая. В прямых проекциях формулы даются для поверхности эллипсоида или шара (в зависимости от масштаба карт), в косых и поперечных - только для поверхности шара.

**Цилиндрические проекции** (cylindrical projection) образуются на основе проектирования земной поверхности на боковую поверхность цилиндра, определенным образом ориентированного относительно земного шара (эллипсоида).

В прямых цилиндрических проекциях параллели и меридианы изображаются двумя семействами параллельных прямых линий, перпендикулярных друг другу. Таким образом задается прямоугольная сетка цилиндрических проекций.

Промежутки между параллелями пропорциональны разностям долгот. Промежутки между меридианами определяются принятым характером изображения или способом проектирования точек земной поверхности на боковую поверхность цилиндра. Из определения проекций следует, что их сетка меридианов и параллелей ортогональна. Цилиндрические проекции можно рассматривать как частный случай конических при  $\alpha=0$  (вершина конуса в бесконечности).

По свойствам изображения проекции могут быть равноугольными, равновеликими и произвольными. Применяются прямые, косые и поперечные цилиндрические проекции в зависимости от расположения изображаемой области. В косых и поперечных проекциях меридианы и параллели изображаются различными кривыми, но средний меридиан проекции, на котором располагается полюс косой системы, всегда прямой.

Существуют разные способы образования цилиндрических проекций. Наглядным представляется проектирование земной поверхности на боковую поверхность цилиндра, которая затем разворачивается на плоскости [3]. Цилиндр может быть касательным к земному шару или секущим его. В первом случае длины сохраняются по экватору, во втором - по двум стандартным параллелям, симметричным относительно экватора.



Цилиндрические проекции применяют при составлении карт мелких и крупных масштабов - от общегеографических до специальных. Так, например, аэронавигационные маршрутные полетные карты чаще всего составляются в косых и поперечных цилиндрических равноугольных проекциях (на шаре).

Существует большое разнообразие проекций: псевдоцилиндрические, поликонические и др. Более подробно об этих проекциях написано в разнообразной специальной литературе.

Программные средства ГИС содержат различные блоки преобразования, включающие следующие виды проекций: равновеликая коническая Алберса, азимутальная равнопромежуточная, коническая равнопромежуточная, квадратная, гномоническая, равновеликая азимутальная Ламберта, коническая равноугольная Ламберта, Меркатора, цилиндрическая Миллера, косая Меркатора (Hotine), ортографическая, полярная стереографическая, поликоническая, синусоидальная, стереографическая, поперечная Меркатора, универсальная поперечная Меркатора (UTM), Гринтена.

При определении проекций исходной карты рекомендуют пользоваться атласом для отечественных карт [13] и для зарубежных [128].

Для топографических карт в отличие от мелкомасштабных нет такого разнообразия проекций, однако имеется разнообразие моделей Земли и географических систем координат.

Специфика российского рынка геоинформационных технологий определяет проблемы проекционных преобразований в России. Одна из серьезных проблем связана с использованием отечественной картографической информации, имеющей значительные отличия от аналогичной иностранной. Как правило, иностранные программные средства не поддерживают напрямую распространенные в нашей стране проекции, а информацию о типе проекции и ее параметрах получить довольно сложно [9].

Следует отметить, что наличие программного обеспечения, осуществляющего проекционные преобразования, отличает ГИС от систем автоматизированного проектирования. САПР могут решать задачи проектирования на локальных участках земной поверхности, для которых можно пренебречь кривизной поверхности.

Кроме того САПР не обеспечивают эффективную стыковку картографической информации разных масштабов и разных проекций, что могут ГИС. Поэтому применение пакетов САПР для реше-

ния задач ГИС является анахронизмом и тормозит развитие геоинформатики.

## 12.6. Цифровые карты

В автоматизированных системах пространственной обработки данных имеется несколько основных типов цифровых моделей: цифровая модель местности, цифровая модель объекта (ЦМО), цифровая модель карты или цифровая карта [109].

Между ЦММ и ЦМК существует различие, которое заключается в том, что ЦММ определяет модель базы данных, ЦМК - модель представления данных, т.е. представление ЦММ.

Можно определить *цифровую карту* как цифровую модель геоинформационной системы, представленную в виде композиции из одного или нескольких слоев. На цифровой карте фиксируются пространственные объекты, связи и отношения между ними, а также пользовательские идентификаторы пространственных объектов, обеспечивающие связь с их атрибутивными данными.

В дополнение к термину "цифровая карта" существует понятие электронной карты. Электронная карта - это отображение цифровой карты с помощью визуальных средств и компьютерных технологий. По информативности электронная карта может быть намного менее содержательной, чем цифровая, на основе которой она построена. Это может быть обусловлено в первую очередь решаемыми задачами. Например, при навигационных задачах (электронная карта в легковом автомобиле) на ней показаны основные факторы способствующие или препятствующие передвижению, а остальные исключаются.

Атрибутивные данные объектов хранятся в виде таблиц, каждая запись в которых соотносится с определенным пространственным объектом цифровой карты через пользовательский идентификатор, указанный и в записи, и в цифровой карте.

Кроме атрибутивного описания содержательная определенность объектов фиксируется в виде конкретных тематических слоев согласно принятой схеме выделения на исходной карте.

Пространственная определенность объектов на цифровой карте фиксируется в соответствующем выделении слоев цифровых карт по типу пространственных объектов (полигоны, линии и точки).

Цифровую карту можно определить как подмножество ЦММ, сформированное для визуального отображения пространственно-временных данных. Следует подчеркнуть, что ЦММ и ЦМК являют-

ся дискретными моделями. Но в процессе отображения ЦМК преобразуется в аналоговую модель - *аналоговую карту*.

## 12.7. Характеристики цифровых моделей

Цифровые модели оперируют с различными типами информации.

**Метрическая информация** передает измерительные характеристики объекта, т.е. координаты, размеры. Эта информация относительно проста и однородна по структуре, в силу чего она является *сильно типизированной*. Метрическая информация в ГИС содержит координатные данные и некоторые (числовые) атрибутивные данные.

Качественным отличием цифровых моделей, полученных по реальным измерениям, является точностная характеристика модели. Она обусловлена ошибками измерений и последующими ошибками вычислений при геометрическом моделировании. Этот параметр определяет применимость цифровой модели в частности, при получении графических реализаций в разных масштабах.

**Атрибутивная информация** - информация о свойствах и связях объектов. В ГИС такая информация включает *атрибутивные данные и метаданные*. Она может подразделяться на семантическую, технологическую и другие виды. Тем не менее все эти виды информации можно назвать семантическими, хотя с учетом сложившейся в ГИС терминологии более правильно называть ее атрибутивной.

Этот тип информации определяет принадлежность точек или объектов к определенному классу или объекту (сложный или простой объект), описывает свойства объектов и их частей, задает взаимосвязи и условия обработки, условия воспроизведения и т.п.

В исходном виде эта информация неоднородна, сложна по структуре и является *слабо типизированной*, поэтому для создания ЦММ требуются анализ, классификация и типизация атрибутивной информации.

Для хранения в базах данных эта информация должна быть типизирована, то есть разбита на более мелкие группы, имеющие достаточное число сходных признаков.

**Синтаксическая информация** - определяет последовательность работы при корректировке или обновлении ЦММ, правила построения и представления ЦММ. Эта информация *типизирована*.

Первые два типа информации определяются логической структурой ЦММ и не зависят от выбора СУБД. В силу этого их можно назвать внутренними по отношению к ЦММ.

Синтаксическая информация зависит от технологии использования ЦММ с учетом конкретных технических средств и их возможностей. Например, она определяется разрешающей способностью монитора и его цветовой палитры, разрешением принтера или плоттера, погрешностью дигитализации или фотограмметрической обработки и т.д. Синтаксическая информация является внешней по отношению к ЦММ, зависит от выбора технических и программных средств.

**Логическая и физическая структура ЦММ.** Понятия логической и физической структуры ЦММ являются развитием и расширением понятий логической и физической модели данных.

Логическая структура ЦММ определяется как совокупность схем и логических записей, описывающих данную ЦММ. Такая характеристика относится к описательным. Схемы, составляющие логическую структуру ЦММ, могут быть различными в зависимости от назначения и принципов построения.

Логическая структура обуславливается концепцией и методологией моделирования. Она может включать схемы взаимосвязи частей ЦММ, соответствующих реальным связям между частями объекта на местности, в базе данных, схемы взаимосвязи свойств ЦММ и схемы построения ЦММ. Она содержит логические записи, составляющие информационную основу. Элементом логической структуры ЦММ является логическая запись.

Физическая структура ЦММ определяется способом реализации логической ЦММ на конкретной технической основе. В частности, она задается форматом записи данных, хранимых на носителях информации. Элементом физической структуры ЦММ является физическая запись.

**Целостность ЦММ.** Недостаточно того, чтобы ЦММ просто содержала набор данных. Необходимо, чтобы она отражала объекты реального мира. Важно, чтобы такое отражение было однозначным и непротиворечивым. В этом случае говорят, что ЦММ удовлетворяет условию целостности (*integrity*). Целостность ЦММ имеет два значения: как объекта БД и как модели реального объекта.

Целостность ЦММ как *объекта базы данных* определяется требованиями СУБД и соответствует понятию целостности информации в БД. Для достижения целостности исходная информа-

ция должна быть типизирована и структурирована. Такая целостность позволяет осуществлять работу с ЦММ как с элементом базы данных, направлять к ней запросы, проводить фильтрацию, получать справки или отчеты.

Целостность ЦММ как *модели реального объекта* определяется требованием получения проекта карты или картографической композиции средствами ГИС. Для достижения такой целостности информация должна быть полной, актуальной и отвечать требованиям точности при получении данного проекта карты. Информация, таким образом, должна включать не только собранные на местности данные, но и библиотеки условных знаков, которые хранятся в БД независимо от ЦММ. В данном случае целостность ЦММ как модели объекта обуславливается полнотой информации БД.

**Дискретность.** ЦММ относится к классу дискретных моделей. Отсюда следует, что ЦММ необходимо хранить как *объекта дискретной базы данных*. Геометрическая часть ЦММ содержит отдельные точки поверхности объектов. Несмотря на это ЦММ позволяет строить непрерывные линии и поверхности, т.е. получать аналоговые модели (аналоговые карты) за счет совместного использования метрической и семантической информации.

Отметим противоречие, заключающееся, с одной стороны, в необходимости выделения большей информативности модели, что увеличивает объем модели, с другой - в необходимости минимизации информационных объемов, обусловленной ограничениями машинных носителей информации и требованием максимальной скорости обработки данных.

**Многофункциональность.** ЦММ создают с целью адаптивности при решении различных задач. Графическое отображение ЦММ не должно зависеть от средств воспроизведения графической информации. Например, одна и та же ЦММ может использоваться для получения карт масштабного ряда.

Для многократного использования ЦММ нужны дополнительные данные: описатели, классификаторы, нормативные данные, правила применения и т.д. Обычно их называют обычно метаданными. Они хранятся в словаре данных (data dictionary).

## 12.8. Виды моделирования

Рассмотрим различные виды моделирования при: сборе и первичной обработке информации, хранении и обновлении, представлении (отображение).

При *сборе информации* для построения цифровых моделей используются автоматизированные средства регистрации и автоматизированных технологий. Источниками информации служат карты, таблицы, спецификации, геодезические координаты точек и объектов местности, координаты точек на аэрокосмических и наземных фотоснимках, данные, получаемые по телевизионным и/или радиолокационным снимкам, телеметрические данные, информация, считываемая с планов и карт, данные о допусках и погрешностях, дополнительная информация текстового характера.

После сбора первичных данных на уровне *хранения и обновления информации* осуществляются симплификация, унификация, коррекция информации, содержащей ошибки и дополнения к ней. Таким образом, формируется унифицированная совокупность данных, одинаковая для различных средств и технологий сбора, позволяющая в дальнейшем применять ее для получения чертежей и планов не одного, а нескольких смежных масштабов.

На уровне *представления ЦММ* цифровая информация отображается в удобном для пользования виде. ЦММ может генерироваться из разных моделей. Визуальное представление ЦММ реализуется на современных устройствах вывода информации.

Технологически можно выделить следующие виды моделирования [98, 105]: семантическое, инвариантное, геометрическое, эвристическое, информационное. Они проявляются в разной степени на разных системных уровнях обработки информации.

*Семантическое* моделирование взаимосвязано с задачами кодирования и лингвистического обеспечения. поэтому оно используется в основном на уровне сбора первичной информации. Это обусловлено также большим объемом и разнообразием входной информации, сложностью ее структуры, возможным наличием ошибок.

Чем более разнородна входная информация по структуре и содержанию, чем менее она унифицирована, тем больший объем семантического моделирования применяется в подсистеме сбора.

*Инвариантное* моделирование основано на работе с полностью или частично унифицированными информационными

элементами или структурами. Его эффективность доказана опытом применения прежде всего САПР и других АС. Этот вид моделирования предполагает использование групповых операций, что обеспечивает более высокую производительность труда по сравнению с индивидуальным моделированием.

Кроме того, инвариантное моделирование способствует повышению производительности обработки информации, особенно при моделировании (обработке) графических объектов.

Примером инвариантного моделирования является работа со слоями в ГИС. Изменение характеристик слоев, влечет за собой изменение всех объектов данного слоя. Одна операция со слоем эквивалентна множеству операций с каждым объектом слоя.

*Геометрическое* моделирование заключается в изменении геометрических характеристики объектов или в построении объектов. Оно применяется там, где требуется обработка метрических данных.

*Эвристическое* заключается в интерактивной обработке. Оно основано на экспертном оценивании результатов каждого шага обработки и последующем принятии решения о выборе очередного метода обработки и этапа обработки. Эвристическое моделирование позволяет учитывать индивидуальные свойства нетиповых объектов и при решении специальных нетиповых задач.

Эвристическое моделирование базируется на реализации общения пользователя с ЭВМ по сценарию, учитывающему, с одной стороны, технологические особенности программного обеспечения, с другой - особенности и опыт обработки данной категории объектов.

*Информационное* моделирование связано с созданием и построением информационных моделей [60]. Оно включает преобразование различных форм информации, например графической или текстовой, в вид, задаваемый пользователем. Это моделирование эффективно только при предварительной разработке интегрированной информационной основы и применении баз данных.

Следует отметить, что степень использования перечисленных видов моделирования в геоинформационных технологиях различна и зависит от этапа обработки информации.

Описание цифровой модели динамично. Оно изменяется или дополняется по мере появления новых задач, новых методов обра-

ботки и новых технических средств автоматизации проектирования.

## **12.9. Цифровые модели рельефа**

**Цифровая модель рельефа** - это дискретная модель рельефа местности (поверхности), предназначенная для компьютерного моделирования и отображения средствами компьютерной графики.

Каждую точку поверхности Земли невозможно передать в модель. Поэтому используют цифровые модели рельефа, которые передают отдельные точки. В тоже время наборы этих точек позволяют восстанавливать всю поверхность при использовании методов компьютерного моделирования.

Главная цель ЦМР - при минимальном числе точек модели обеспечить максимально возможную точность построения рельефа.

В простейшем случае ЦМР это набор трехмерных координат точек рельефа и информация о связях между точками и способах восстановления поверхности по данным точкам.

Точки для ЦМР могут собираться по горизонталям, по профилям и по заданной регулярной или нерегулярной сетке.

Сетку задают плановые координаты. Регулярная сетка предпочтительнее нерегулярной тем, что она существенно (на порядки) снижает размерность матриц, необходимых при аналитическом описании и восстановлении любой точки рельефа.

Основная проблема при построении ЦМР заключается в том, что реальная поверхность является нерегулярной, т.е. имеет разрывы. Математические модели, как правило, описывают гладкую, регулярную поверхность.

Поэтому довольно часто цифровую модель рельефа строят из совокупностей разных математических моделей, стыкующихся в определенных точках. Одним из таких подходов является метод Вороного-Делоне.

## **12.10. Построение цифровых моделей рельефа на основе метода Вороного-Делоне**

### **12.10.1. Историческая справка**

Георгий Феодосеевич Вороной родился 16 (28) апреля 1868 г. в селении Журавка Полтавской губернии. Закончив гимназию в Прилуках, с 1885 по 1890 гг. Георгий учился в Санкт-Петербургском университете, где связал свою судьбу с математикой.



После защиты магистерской диссертации в 1894 г. он получает направление в Варшавский университет. В 1898 г. Вороной становится профессором Варшавского политехнического института. В 1903 – 1904 гг. вышли его две большие статьи по теории чисел, внесшие существенный вклад в эту область, в результате чего в 1907 г. он был избран членом-корреспондентом Санкт-Петербургской академии наук.

Арифметическая теория положительно-определенных квадратичных форм представляет одно из самых значительных достижений Вороного, обессмертивших его имя. Во время революции 1905 – 1907 гг. Вороной переезжает в Новочеркасск. Там обострилась его желчекаменная болезнь, из-за которой он умер в ноябре 1908 г. Похоронен Г.Ф. Вороной согласно его завещанию на его родине – в Журавке.

В 1908 г. вышла первая ("Свойства положительных совершенных квадратичных форм"), а в 1909 – вторая часть ("Исследования о примитивных параллелоэдрах") работы Вороного о квадратичных формах. Задача, которую рассматривает Вороной, посвящена проблеме заполнения  $n$ -мерного евклидова пространства параллелоэдрами, не входящими друг в друга со смежными гранями.

В то время не стояла проблема построения модели рельефа. Но созданная Г.Ф. Вороным почти сто пятьдесят лет назад модель в настоящее время применяется для описания рельефа многими специалистами без упоминания имени его создателя.

Борис Николаевич Делоне родился 3 (15) марта 1890 г. в Санкт-Петербурге. Однако вскоре семья переехала в Киев, где его отец получил должность профессора механики. С 1908 по 1913 гг. Борис Николаевич учился на физико - математическом факультете Киевского университета и был оставлен при университете "для подготовки к профессорскому званию". До 1922 г. он преподавал в Киевском университете, а затем переехал в Ленинград, где стал профессором Ленинградского университета.

В 1929 г. Делоне избирают членом-корреспондентом АН СССР, а в 1935 г. он переезжает в Москву, куда переводилась Академия наук. С тех пор 45 лет, он жил в Москве, работал в Математическом институте им. В.А. Стеклова и преподавал в Московском университете на кафедре геометрии. Умер Борис Николаевич в 1980 г., в возрасте 90 лет, прожив долгую и плодотворную жизнь, внося свой вклад в теорию чисел, алгебру и геометрию, а также воспитав многих выдающихся учеников.

### 12.10.2. Метод Вороного-Делоне

Одна из проблем моделирования рельефа заключается в нахождении способа, который позволял бы при минимальном количестве точек модели максимально информативно отображать рельеф, т.е. исходную поверхность.

Подход, дающий решение данной проблемы - метод Вороного-Делоне. Метод упрощенно можно свести к нахождению системы дискретных точек, заполняющих некое пространство. Первоначально работы были выполнены Вороным, затем их развил и обобщил Делоне.

Г.Ф. Вороной предложил разбиение пространства, содержащего точки (центры) на участки, в каждом из которых имелась точка, служащая основой построения многогранника, заполняющего пространство.

Разбиение Вороного включает понятия: плоскость, канал и узел - Вороного. Для двух произвольных центров (точек) плоскость Вороного есть геометрическое место точек, равно удаленных от этих двух центров.

Канал Вороного строится для трех центров (треугольник). Канал Вороного есть геометрическое место точек, равно удаленных от этих трех центров.

Узел Вороного строится для четырех центров (тетраэдр). Узел Вороного есть геометрическое место точек, равно удаленных от этих четырех центров.

При построении рельефа на нем могут быть выбраны произвольные точки, которые отражают характерные точки рельефа. В плане эти точки могут образовывать сеть многоугольников.

На рис. 12.1. показан многогранник Вороного. Для точки  $i$  и окружающих ее точек 1-5 строятся плоскости Вороного. Пересечения этих плоскостей образует фигуру называемую многогранником Вороного. Внутри данного многогранника все точки пространства ближе к точке  $i$ , чем к любому другому центру. Следовательно точка  $i$  может служить центром дискретизации пространства, заключенного в многограннике Вороного. Она может служить точкой цифровой модели, аппроксимирующей часть пространства.

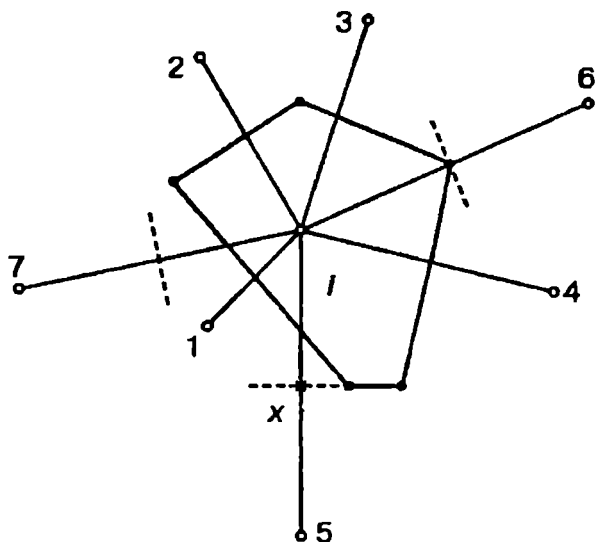


Рис.12.1 Многогранник Вороного

Множество многогранников Вороного образует набор точек оптимально дискретизирующих произвольную поверхность в плане.

Делоне развил эту теорию и рассмотрел случай, когда совокупность точек (центров) можно представить как набор треугольников. Очевидно, что через каждую тройку точек можно построить описанную окружность (рис.12.2.). Центром этой окружности будет точка, лежащая на канале Вороного.

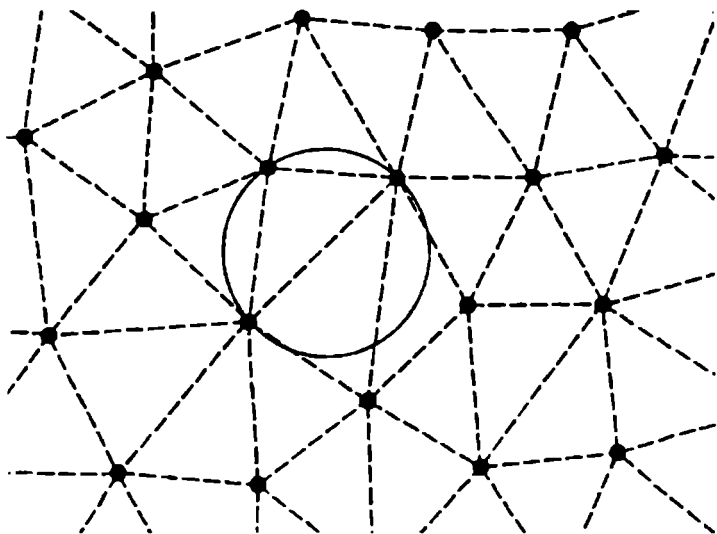


Рис. 12.2. Разбиение Делоне

Предположим, что плоские треугольники рис.12.2. являются проекцией поверхности. Эта поверхность пересекает канал Вороного и точка пересечения имеет отметку высоты  $z$ . На этих четырех точках можно построить тетраэдр, который называют симплексом Делоне. Таким образом, можно построить совокупность тетраэдров, которые заполнят пространство без щелей и разрывов. Эта совокупность называется разбиением Делоне.

Такое разбиение позволяет создать пространственную треугольную сеть оптимальным образом описывающую произвольную поверхность. Такой алгоритм может служить основой для построения цифровой модели рельефа.

### 13. Геоинформационное моделирование

Как уже говорилось, в геоинформатике применяются разные виды моделирования: эвристическое, информационное, геометрическое и др. Это обусловлено интеграцией в геоинформационные технологии технологий других автоматизированных систем. Эти виды моделирования могут использоваться в любых информационных системах.

В геоинформатике есть специфическое моделирование, которое применяется только в геоинформационных системах. Это геоинформационное моделирование [74, 99].

Геоинформационное моделирование это преобразование моделей пространственных объектов, использующее их организацию, согласно которой каждый графический объект взаимосвязан с одной или несколькими таблицами базы данных, а преобразование графического объекта влечет преобразование табличных данных. При этом преобразование табличных данных меняет характеристики графических объектов.

Геоинформационное моделирование определяют также как класс моделирования графических объектов, взаимосвязанных с базами данных и включающими пять основных типов моделирования. Перечислим эти типы:

1. Преобразование графической информации, которое приводит к изменению графических и табличных данных;
2. Преобразование табличных данных, что приводит к изменению графических и табличных данных;
3. Преобразование графических объектов из одного типа в другой;
4. Построение цифровых моделей явлений;
5. Построение, редактирование или модификация графических объектов на основе отношений между пространственными объектами (без использования графических редакторов).

Основу геоинформационного моделирования как специализированной технологии составляют преобразования основанные на теоретико-множественных отношениях, законах формальной логики, алгоритмах обработки изображений, технологиях работы с компьютерной графикой, технологиях СУБД и многом другом.

Объектами геонформационного моделирования являются пространственные графические объекты и объекты базы данных ГИС.

Геонформационное моделирование включает следующие специальные технологии:

*геогруппировку* - построение временной динамической модели путем объединения графических объектов в более крупные объекты;

*буферизацию* - процедуру построения полигональных объектов по заданным линейным (и точечным) объектам и параметрам буферизации;

*генерализацию* - процедуру обобщения графических объектов и изменения их видимости при изменении масштаба;

*комбинирование* - процедуры композиции или декомпозиции графических объектов на основе отношений между ними;

*геокодирование* - процедуру позиционирования (координатной привязки) данных одной таблицы к данным другой, позиционно определенной таблицы;

*обобщение данных* - процедуру создания атрибутов новых объектов на основе отношений атрибутов исходных объектов;

*построение тематических карт* на основе анализа и обработки атрибутивных данных;

*проведение автоматической классификации* признаков графических объектов (включая растровые) по заданным критериям.

### 13.1. Формализация технологических процессов

При геонформационном моделировании реальное пространственное явление или объект описываются с помощью специального аппарата. Такое описание называют *формализованным*. Оно служит для представления исследуемых элементов явлений и их взаимосвязей.

Геонформационное моделирование является технологическим. В отличие от теоретических методов моделирования оно взаимодействует с объектами базы данных ГИС.

Процесс геонформационного моделирования оператор начинает с того, что выделяет объект или объекты моделирования среди множества других, не участвующих в процессе моделирования.

Данная процедура называется *активизацией* объекта (объектов) и осуществляется по аналогии с процедурой активизации объектов во всех существующих графических редакторах.

Будем обозначать активизированный объект символом  $A$ , прочие объекты символом  $O$ .

При геоинформационном моделировании могут использовать специальный объект  $S$ , который применяют в качестве шаблона (маски).

Объект вновь создаваемый или модифицируемый из изменяемого объекта обозначим символом  $M$ .

В рамках такой модели обобщенная процедура геоинформационного моделирования описывается на основе отношений между  $A$  и  $S$

$$A \times S \rightarrow M$$

Где

$\times$  - символ отношения;

$\rightarrow$  импликация

В геоинформатике имеются три типа графических векторных объектов: точечные ( $Pt$ ), линейные ( $Ln$ ) и площадные ( $Ar$ ). Изменяемый объект  $A$  и  $S$  шаблон должны принадлежать к одному типу.

Условимся обозначать совокупность атрибутов произвольно-го объекта  $O$  символом  $At(O)$ , совокупность позиционных данных символом  $Pos(O)$ .

### 13.2. Комбинирование выбранных объектов

При таком виде моделирования осуществляют операции с графическими объектами.

Активизированный объект  $A$  называют также изменяемым объектом, подчеркивая этим саму возможность его модификации и взаимодействия с другими подобными объектами. Однако из этого не следует, что этот объект надо в обязательном порядке изменять.

**Объединение объектов.** Наиболее распространенной процедурой комбинирования объектов является их объединение. Эту процедуру удобно использовать в тех случаях, когда один из объектов является определяющим (например, материковую часть государства и острова, относящиеся к нему, в одну территорию).

$$A1 \cup A2 \cup A_i \cup A_k \rightarrow M$$

$k$  - число исходных объектов, участвующих в объединении.

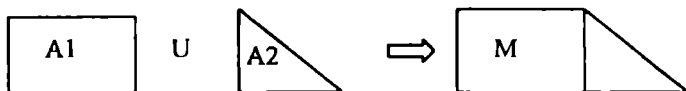


Рис. 13.1. Комбинирование объектов на основе объединения

В результате такого вида моделирования создается новый объект и вычисляются соответствующие ему данные.

**Выделение нескольких объектов из одного.** Процедуры выделения объектов позволяют разбить изменяемый объект на более мелкие объекты, используя некоторый шаблон. Например, с их помощью можно разбить большую территорию на составляющие части.

$$A \cap S1 \rightarrow M1 ; A \cap S2 \rightarrow M2 ; \dots A \cap Sn \rightarrow Mn.$$

**Удаление фрагментов объектов путем комбинирования.**

Данную операцию осуществляют также как операцию моделирования с несколькими объектами. Удалить часть изменяемого объекта можно путем наложения на него объекта-шаблона.

$$A - \neg A \cap S \rightarrow M$$

Другой путь - удаление фрагмента изменяемого объекта, который не совпадает с объектом-шаблоном.

$$A \cap S \rightarrow M$$

Например, для того, чтобы удалить из объекта, обозначающего границы административной единицы, территорию некоторого озера, надо создать озеро как шаблон и с помощью команд ГИС удалить фрагмент административной единицы, налегающий на озеро.

**Добавление узлов путем комбинирования объектов.** Процедура моделирования "Добавление узлов" позволяет создать новые точки в полигоне с сохранением топологии или с ее изменением, если добавляемый узел соединяет линейный объект с другим.

В отличие от обычной процедуры графического редактирования добавление узлов - это процедура геоинформационного моделирования, основанная на взаимодействии не менее двух линейных объектов. Само добавление узлов осуществляется автоматически, как результат суперпозиции графических объектов, а не с помощью указаний оператора как в технологиях компьютерной графики.

$$\exists (A1, A2) : Pos(A1) \wedge Pos(A2) \rightarrow Pos(M)$$

Подобная процедура нужна, например, при внесении на карту новой улицы, соприкасающейся с уже существующими улицами. Она позволяет точно задать места пересечения улиц в виде узлов.

### 13.3. Преобразование атрибутов комбинируемых объектов

При создании нового объекта необходимо определить его атрибуты, т.е. сопоставить ему таблицу с данными. Объект может созда-



ваться как путем укрупнения, так и на основе разбиения на части более крупного объекта.

При объединении (композиции) более мелких, имеющих табличные данные, необходимо задать правила вычисления атрибутов аналогичных данных для создаваемого нового объекта. Такие процедуры геоинформационного моделирования называют обобщением данных исходных объектов.

В процессе обобщения вычисляются данные для создающегося объекта в зависимости от заданного метода отношений атрибутов. Большинство ГИС - технологий содержит следующие методы обобщения данных при объединении объектов:

**сумма** - значения атрибутов, соответствующих исходным объектам, складываются, и сумма присваивается новому объекту.

$$At(M) = \sum At(A_i) \quad i=1.....k$$

$k$  - число исходных объектов.

**среднее** - вычисляется среднее значение атрибутов исходных объектов и присваивается атрибуту нового объекта.

$$At(M) = \sum At(A_i) / k \quad i=1.....k$$

**взвешенное среднее** - разные значения для исходных объектов умножаются на различные коэффициенты (веса)  $p$ . Веса можно брать из любого числового поля таблицы или вычислять по значению характеристики пространственного объекта (например, использовать его площадь или периметр, которые могут отсутствовать в таблице.)

$$At(M) = \sum p At(A_i) / k \quad i=1.....k$$

**значение** - атрибуту нового объекта присваивается указанное значение  $Z$ .

$$At(M) = Z$$

**без изменений** - атрибуту нового объекта присваивается значение, соответствовавшее исходному объекту.

$$At(M) = At(A)$$

При декомпозиции пространственных объектов используют следующие методы преобразования атрибутов:

**пусто** - удаляет значение, которое соответствовало изменяемому объекту.

**значение** - сохраняет значение, которое соответствовало изменяемому объекту.

**пропорционально размеру** - вычитает из значения (которое соответствовало изменяемому объекту)

долю, пропорциональную размеру вырезанного фрагмента.

Рассмотренные процедуры легко реализуются в системах баз данных, электронных таблицах, пакетах статистической обработки в виде стандартных функций.

#### **13.4. Геоинформационное моделирование с использованием геогрупп**

Технология моделирования, называемая геогруппировкой, основана на известном во многих пакетах компьютерной графики процессе группировки (объединения с сохранением связей) графических объектов. Отличие заключается в некоторых дополнительных возможностях, которые геогруппировка дает пользователю.

*Во-первых*, это возможность получения на основе объединения не одного, а **нескольких графических объектов**. Эта часть моделирования, по большому счету, не имеет никакого отношения ни к географической информации, ни к термину "гео" и является простым расширением технологии группировки графических объектов в компьютерной графике.

*Во-вторых*, процесс геогруппировки наряду с объединением графических объектов сопровождается **объединением их атрибутов**, хранящихся в базе данных, связанной с графическими объектами.

Общий класс моделирования, частным случаем которого является геогруппирование, это процесс построения визуальных динамических моделей совокупностей графических объектов с объединением их характеристик, хранящихся в связанной с графическими объектами базе данных.

Целью геогруппировки является построение новых графических объектов, называемых **геогруппами**.

Процесс построения геогрупп состоит в объединении пространственных объектов в группы в соответствии с заданными признаками.

Таким образом, данный вид моделирования создает условную (тематическую) карту методом индивидуальных значений, в которой тематической переменной является территория нового сгруппированного объекта. Особенностью такой тематической карты является ее *временный* характер. Ее необходимо анализировать и в случае необходимости зафиксировать с помощью специальных процедур.

Исходя из вышеизложенного можно сформулировать следующий вывод:

технология моделирования, основанная на построении гео групп, представляет собой процесс построения динамической модели графических данных и их атрибутов как средства анализа и поддержки принятия решений.

### 13.5. Построение новых графических объектов на основе слияния атрибутивных данных

Атрибутивные данные хранятся в табличной форме в базе данных ГИС. Поэтому построение новых объектов на основе слияния их атрибутов использует возможности интерфейса ГИС по работе с табличной информацией, в частности на использовании команд типа "Слияние в таблице". Формализовано данная процедура может быть описана секвенцией вида

$$At(A1), \dots, At(An) \Rightarrow A1, \dots, An$$

и отношением вида

$$A1 \cup A2 \cup Ai \cup Ak \rightarrow M$$

Процедуры слияния атрибутивных данных в таблице позволяют создавать новые графические объекты путем объединения существующих объектов подобно процессу гео группировки. Атрибутивные данные нового объекта, полученного таким способом, вычисляются на основе процедур обобщения. Отличие данной процедуры от гео группировки технологическое. В данном случае соединение объектов происходит по табличным данным без использования графических объектов.

В процедуре гео группировки соединение происходит по графическим объектам без привлечения табличных данных. В обоих случаях создается временная модель, которая при необходимости может быть зафиксирована.

### 13.6. Буферизация

Буфером или буферной зоной называют область, которая охватывает все объекты, расположенные не далее заданного расстояния от некоторого линейного объекта, области, символа или иного объекта в окне. Процесс буферизации означает построение цифровой модели некоего явления с помощью процедур геоинформационного моделирования.

При буферизации на основе объектов одного типа создаются объекты другого типа. Так линейный или точечный объекты служат основой для создания площадных объектов.

Вид буфера определяется его радиусом и формой объекта, вокруг которого строится буфер, и способом построения буфера. Радиус буфера задается либо в виде константы, либо в виде колонки таблицы, либо в виде выражения, создаваемого на основе атрибутов таблицы.

В некоторых ГИС для графического представления буфера необходимо указать гладкость (число сегментов для буферной окружности).

Применение значений табличных данных для построения буферной зоны относится к случаю геоинформационного моделирования, использующему построение новых графических объектов на основе обработки неграфических (табличных данных).

Радиус буферной зоны может задаваться в виде значений из столбца таблицы базы данных ГИС. Например, чтобы создать вокруг городов буферные зоны, которые отражали бы численность их населения, можно выбирать значения радиуса буфера из колонки "Население". Такой прием применяется как дополнительная возможность тематического картографирования.

Кроме того, единый буфер можно создавать вокруг всех выбранных объектов или отдельные буферные зоны вокруг каждого объекта. На Рис. 13.2. Показаны разные типы буферных зон.

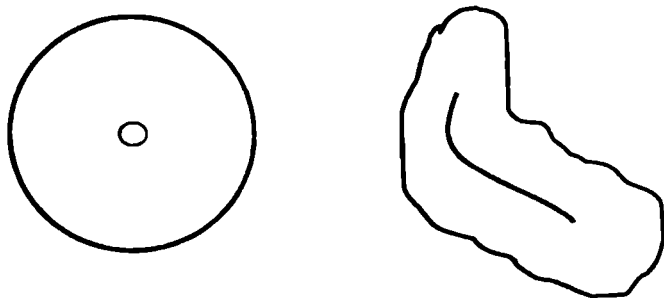


Рис. 13.2. Буферные зоны

### 13.7. Геокодирование

Процессом позиционирования называют пространственную привязку табличных данных к точкам земной поверхности или к

системе координат карты. Позиционирование является обязательным свойством данных в геоинформатике.

Атрибутивные данные могут собирать независимо от геоинформационных технологий. Например, статистические данные о субъектах федерации собираются в Госкомстате. Эти данные необходимо привязать к соответствующим изображениям этих субъектов на карте.

Для этой цели используют процедуру геокодирования. При этом необходимо, чтобы имелась хотя бы одна таблица (например с кадастровыми номерами субъектов федерации), связанная с картографическими отображениями этих субъектов.

Процедура геокодирования позволяет связывать таблицы, у которых есть хотя бы один столбец с одинаковыми атрибутивными данными. Такая связь позволяет объединять вновь собранные данные с графическим отображением свойств пространственных объектов.

В результате геокодирования таблица с новыми данными становится позиционированной, т.е. ее данные получают пространственную привязку.

Поскольку сравнение данных таблиц можно проводить разными путями, существуют разные способы геокодирования.

При геокодировании по *полному адресу* сравниваются адреса в кодируемой таблице с информацией об улицах и адресах в таблице специального формата, которая поставляется как дополнительное приложение к конкретной ГИС.

Геокодирование таблицы *по областям* (площадным объектам) основано на сравнении названия области в записи из кодируемой таблицы с названиями областей в таблице поиска. В этом случае записи таблицы присваивает координаты  $X$  и  $Y$  центроида соответствующей области из таблицы поиска. Центроид области – это примерная точка центра области (координаты центра описанного вокруг области прямоугольника).

Помимо описанных выше способов, существует грубое геокодирование. Оно заключается не в точном определении местоположения объекта (например, адреса), а в указании района его расположения. Для этого достаточно использовать в качестве таблицы поиска файл, содержащий координаты центроидов районов. Следует отметить, что по этой причине в ГИС предусмотрен режим ручного геокодирования. Во многих методах геокодирования могут возникнуть ситуации, в которых нельзя добиться полного совпадения.

Результаты геокодирования могут быть записаны в поле таблицы, в виде числа, показывающего, какие действия по сравнению были произведены или почему поиск не удался. С помощью кода результата можно находить разные типы необработанных записей.

Процесс геокодирования позволяет организовать поиск местоположения объекта на графическом изображении карты по заданным атрибутам. Найденный объект отмечается символом, определенным заранее в системных параметрах ГИС.

### 13.8. Классификация

Современные ГИС широко используют не только картографическую и векторную информацию, но и данные фотосъемок и соответственно растровые изображения. Векторные данные называют объектными, растровые – полевыми. Последние широко используются для изучения и выявления новых явлений.

Одним из методов геоинформационного моделирования, позволяющим выявлять новые объекты, является процедура *классификации* данных. Она основана на том, что все пиксели изображения, имеющие сходные спектральные сигнатуры, объединяются в класс.

Эта процедура полностью заимствована из методов распознавания образов и является результатом интеграции технологий обработки данных дистанционного зондирования (ДДЗ) в ГИС-технологии.

**Классификация** заключается в определении принадлежности исследуемого объекта к известному классу. Она производится автоматически по набору признаков, которые задает оператор.

Классификацию можно проводить по одному или нескольким классам, с обучением и без него. Она может быть использована для последующей векторизации изображений или улучшения качества изображения.

## 14. Основные технологические этапы обработки данных в ГИС

Несмотря на разнообразие программных реализаций геоинформационных систем можно выделить общие подходы и общие принципы в работе любой ГИС.

В качестве примеров технологических подходов в отдельных случаях будут использованы возможности комплекса программных средств MapInfo как одного из наиболее простых пакетов.

Рассмотрим технологию векторных ГИС. Основными объектами, с которыми оперирует такая геоинформационная система, являются векторные объекты.

Построение карты в ГИС заключается в создании совокупности векторных объектов, отображающих необходимые объекты местности и привязки векторных данных к таблицам, в которых хранится атрибутивная информация.

Основные технологические процессы в векторной ГИС делятся на следующие группы: ввод данных, векторизация, графическое редактирование, геоинформационное моделирование, подготовка данных к печати. Основы геоинформационного моделирования рассмотрены в главе 9 данного учебника. Ниже рассмотрим остальные составляющие обобщенной ГИС-технологии.

Построение карты может быть реализовано тремя путями:

1. Создание новой карты на основе информации, которая вводится оператором,
2. Создание новой карты на основе существующей векторной карты путем ее модификации или обновления,
3. Создание новой карты на основе векторизации (трассировки) растровых изображений, которые могут представлять собой сканированные снимки или карты.

### 14.1. Основные функции интерфейса ГИС

Создание новой карты на основе информации вводимой оператором с клавиатуры наименее производительное и на практике не применяется. Технологически этот процесс используется при редактировании карт.

При работе с ГИС оператор использует графический интерфейс, который задает основные и вспомогательные режимы работы, позволяет устанавливать настройки и параметры визуализации.

**Графический интерфейс** представляет собой набор кнопок и падающих меню, обеспечивающих работу команд, режимов и функций ГИС. Особенность графического интерфейса в том, что большинство команд падающего меню дублируются специальными кнопками.

Иногда команды падающего меню называют опциями. К числу наиболее общих опций большинства ГИС относятся следующие:

- Добавить выборку в группу,
- Режимы выбора объектов,
- Перенос объекта ,
- Изменить вид изображения на мониторе объекта
- Перемещение изображения на мониторе (pan),
- Вывод внутренней информации о выделенном объекте (info),
- Ввод текста,
- Управление слоями
- Легенда
- Показ линеек изображения
- Одиночный выбор объекта
- Применение функций статистики
- Изменение масштаба изображения (Зуммирование),
- Редактирование узлов
- Редактирование дуг
- Рисование графических примитивов
- Изменение формы графического объекта
- Группирование графических объектов
- Разгруппирование графических объектов
- Создание или изменение стилей (линий, точек, полигонов, текста)
- Создание специальных форм (фреймов)

Этот перечень содержит примерный список опций. ГИС может либо включать в свой состав дополнительные опции, либо не включать что –то из перечисленных в списке.

Каждая инструментальная ГИС содержит главное падающее меню, которое содержит разные разделы, но общими разделами меню, как правило, являются разделы команд: Файл, Правка, Запрос, Таблица, Настройки, Окно, Карта, Справка.



При использовании команд падающего меню название меню будет отделено от команды меню символом > . Например, команда "Открыть карту" с использованием меню "Файл" в такой нотации будет выглядеть как - Файл > Открыть карту.

## 14.2. Подготовительный этап

При работе с растровыми данными возможны два варианта:

Работа с новой картой – создание,

Работа с существующей картой – редактирование или обновление.

В первом случае используют команду Файл > Новая Карта. Во втором команду Файл > Открыть.

Следует сразу обратить внимание на то, что в некоторых ГИС термин «карта» может заменяться другим. Например, в MapInfo вместо термина «карта» используется термин «таблица», а для обозначения таблицы применяется термин «список». И хотя в меню этой ГИС все же есть опция «карта», полностью функции работы с картой определяются разделами меню «таблица» и «карта». В других ГИС используется понятие "проект", что подчеркивает сходство проектирования карт с проектированием чертежей в САПР.

Во многих ГИС требуется создать не один слой карты, а набор слоев, образующих картографическую композицию. С таким набором удобно работать как с единой совокупностью данных.

Реализация этой возможности обеспечивается режимом, который называется «Рабочий набор» или "Проект".

Преимущество использования Рабочего Набора в том, что при его использовании сохраняется список открытых карт, таблиц и окон, а также вспомогательные данные об их расположении на экране в той последовательности, какую они имели в процессе предыдущего сеанса работы.

Открытые ранее таблицы (включая временные) и окна загружаются вновь при открытии Рабочего Набора. Если после загрузки Рабочего Набора открытые ранее окна и таблицы больше не понадобятся, то перед загрузкой Рабочего Набора необходимо выполнить команду «Файл» > «Закреть все». В противном случае таблицы и окна нового рабочего набора добавятся к окнам и таблицам, открытым в течение данного сеанса.

При копировании информации, например на дискету, необходимо копировать не только рабочий набор, но и дополнительно все карты (слои) входящие в его состав. Это обусловлено тем, что файл

рабочего набора всегда содержит информацию о последовательности, структуре и директориях слоев, образующих картографическую композицию, но может не содержать сами файлы

### **14.3. Применение файлов в формате DXF для создания картографической основы**

При работе с конкретной ГИС возможно использование в качестве начальной картоосновы графических материалов из других информационных систем. Для использования карт, созданных в других ГИС или в системах САПР имеющих другие форматы, применяют специальный обменный формат DXF.

Для этого в другой ГИС или иной системе (например, AutoCAD) осуществляют конвертирование файлов из внутреннего формата в формат DXF

Ввод файлов в ГИС осуществляется с помощью процедуры «Импорт» в формате DXF

Импорт в формате DXF использует большое количество проекций и систем координат. В большинстве случаев это определяет необходимость применять преобразование координат.

Преобразование координат необходимо при импорте карты, созданной в системе координат, не используемой в данной ГИС. Преобразование координат необходимо при импорте карты (чертежа), координаты которой не имеют географического смысла, например, если карта создана в обычном графическом редакторе.

При работе с новой картой или после импорта файла необходимо установить единицы измерения точек карты.

## **14.4. Работа с растровыми данными**

### **14.4.1. Ввод растрового изображения**

Исходное растровое изображение представляет собой набор пикселей, не имеющих пространственной привязки. Поэтому на этапе ввода растровые изображения необходимо позиционировать (делать позиционную привязку к системе координат), если оно предназначено для работы с векторными позиционированными данными. Файлы позиционированных растровых изображений содержат контрольные точки привязки, определенные при вводе

Для позиционирования (регистрации) изображения предварительно должны быть определены точки, координаты которых известны.

Позиционирование растрового изображения возможно двумя методами. В первом координаты контрольных точек карты вводятся пользователем с клавиатуры, во – втором – они определяются по существующей векторной карте.

Первый метод регистрации обычно применяется, когда необходимо векторизовать растровое изображение, для которого отсутствует его векторный аналог. Другими словами такая регистрация необходима при создании новой карты. Такой подход требует предварительного определения координат точек карты, которые будут использоваться в качестве контрольных. Следовательно, в распоряжении оператора должна быть некая таблица с координатами возможных контрольных точек.

Эффективность этого метода повышается в тех случаях, когда контрольные точки маркированы, т.е. выделены на растровом изображении и на его бумажном оригинале.

Если растровое изображение получено сканированием карты, то при вводе такого изображения необходимо задать картографическую проекцию, которую имеет используемая карта.

При вводе координат обычно используют то правило, что любая точка, находящаяся западнее нулевого меридиана, имеет отрицательную долготу. Любая точка, находящаяся южнее экватора, имеет отрицательную широту. Так, 52 градуса западнее Гринвича соответствуют X-координате -52.

Для обеспечения более точных результатов позиционирования растрового изображения вводят избыточные измерения. Каждая контрольная точка помогает связывать земные координаты с положением объектов на растровом изображении. В идеале контрольные точки должны задаваться в каждом углу растрового изображения или хотя бы недалеко от каждого угла. Такой способ задания повышает точность привязки растра к векторному изображению.

Число необходимых контрольных точек зависит от типа проекции изображения. Если пользователь не имеет возможности определить проекцию карты или работает с изображением с аэрофотоснимка, то ему может понадобиться большое число контрольных точек.

Второй метод регистрации растровых изображений основан на вводе координат контрольных точек векторного изображения путем их выбора из существующей векторной карты. Технологически эта процедура применяется, когда необходимо привязать растровое изображение к уже существующей в электронном виде векторной карте.

Другими словами такая регистрация необходима при обновлении существующей карты. Данный подход не требует предварительного определения координат точек карты, которые будут использоваться в качестве контрольных. При этом оператору не требуется таблица с координатами возможных контрольных точек.

Однако для этого метода необходимо предварительное опознавание и идентификация контрольных точек на растровом изображении и на его электронном векторном оригинале.

После выполнения регистрации изображения ГИС создаст специальный табличный файл, в котором будет сохранена информация о позиционировании. В следующий раз это растровое изображение можно открыть как таблицу командой "Файл > Открыть карту" без повторения процедур позиционирования.

При позиционировании и определении координат контрольных точек возможны ошибки, которые надо устранить или уменьшить их количество. Это достигается в последующем использовании режима изменения положения контрольных точек.

При позиционировании растрового изображения для контрольных точек выдаются погрешности позиционирования. Величина погрешностей должна быть минимальной. При больших погрешностях могут возникнуть существенные ошибки совмещения векторного и растрового слоев. Наличие больших погрешностей позиционирования требует последующей коррекции.

Точность позиционирования можно повысить, увеличив число контрольных точек или изменив положение контрольных точек.

#### **14.4.2. Векторизация растрового изображения**

Векторные модели используют в качестве основной модели последовательность координат, образующих линию. Векторные модели получают разными способами. Один из наиболее распространенных - векторизация сканированных (растровых) изображений. Она заключается в выделении векторных объектов со сканированного изображения и получении их в векторном формате.

Для векторизации необходимо высокое качество (отчетливые линии и контуры) растровых образов. Чтобы обеспечить требуемую четкость линий иногда приходится заниматься улучшением качества изображения.

Сам процесс сканирования занимает незначительное количество времени, но необходимость последующей векторизации увеличивает расходы практически до уровня ручного цифрования. Сле-

дует отметить, что при векторизации возможны ошибки, исправление которых осуществляется в два этапа:

1) корректировка растрового изображения до его векторизации;

2) корректировка векторных объектов.

Векторные модели с помощью дискретных наборов данных отображают непрерывные объекты или явления. Следовательно, можно говорить о векторной дискретизации. Векторное представление (в отличие от растрового) позволяет отразить большую пространственную изменчивость для одних районов, чем для других, что обусловлено более четким показом границ и их меньшей зависимостью от исходного образа (изображения), чем при растровом отображении. Это типично для социальных, экономических, демографических явлений, изменчивость которых в ряде районов более интенсивна.

При построении новых полилиний или полигонов по растровому изображению в ряде ГИС применяют режимы полуавтоматической (трассировку) или автоматической векторизации для отслеживания существующих линий или полигонов.

Чтобы произвести трассировку (полуавтоматический процесс) узлов полилинии или полигона необходимо активизировать режим трассировки. Выбирают тип трассируемого объекта: Полигон или Полилиния. Затем выделяют границу области трассируемого объекта в начальной точке и запускают режим трассирования.

В случае возникновения сбоя или тупиковых ситуаций повторяют процедуру указания границы объекта.

Трассировка работает только для одного объекта в каждый данный момент времени. При выделении другого объекта будет нарисована прямая линия между двумя узлами этих объектов.

#### **14.5. Редактирование графических данных**

Современный графический интерфейс пользователя ГИС основан на графических стандартах MS Windows. По этой причине современный пользователь геоинформационной системы должен знать операционную систему MS Windows и ее графический интерфейс. В частности, пользователь ГИС должен знать работу простейшего графического редактора.

Считая, что пользователь ГИС достаточно просто выполнить команды рисовки графических примитивов, остановимся на специфической команде «Добавить узлы». При использовании этой ко-

манды Программное средство ГИС добавляет узлы к текущим изменяемым объектам, что находить отражение в изменении соответствующих топологических свойств объектов.

При добавлении узлов к отрезку прямой линии ГИС преобразует линию в ломаную (полилинию). Добавление узла к эллипсу или прямоугольнику превращает этот объект из типа «линейного» в объект типа «область».

Удаление узлов основано на их выделении (одиночном или групповом) и выборе режима удаления. Выбранный узел или узлы будут удалены.

Если пользователь хочет поместить удаленные узлы в буфер обмена, то он может это сделать с помощью команды Правка > Вырезать.

Табличные данные при этой процедуре не копируются. При копировании из буфера обмена помещенных туда узлов пользователь получает объект, имеющий запись с пустыми табличными полями.

В случае когда удаляются все узлы, объект - полилиния или область - будет удален полностью. При удалении несколько узлов - объект перерисовывается.

При удалении узла из объекта типа «полигон» (область) контур полигона не разрывается и объект остается объектом того же типа, т.е. областью. Для преобразования области в полилинию используется команда Объекты > Превратить в полилинии.

В процессе работы может возникнуть необходимость передачи данных через буфер обмена со слоя на слой, с карты на карты, из одного Рабочего набора в другой или из MapInfo в другое приложение Windows. Такая процедура осуществляется через буфер обмена (Clipboard).

Особенностью многих ГИС в отличии от графических редакторов является необходимость установки режимов копирования в буфер для разных типов графических объектов. Такие установки в диалоге "Внешние (системные) режимы" определяют, как происходит обмен информацией ГИС с другими программами через системный Буфер обмена Windows при выполнении команд Вырезать и Копировать.

Флажок "Копировать текст" - Устанавливает режим копирования текста в буфер обмена. Флажок "Копировать изображение (растр)" - Устанавливает режим копирования растра.

Флажок "Копировать картинку (метафайл)" - Устанавливает режим копирования метафайла. Если пользователь не хочет, чтобы

при выполнении команд Вырезать и Копировать в буфер обмена помещались метафайлы, он должен сбросить флажок "Копировать метафайл".

Для копирования и текста, и графики устанавливают оба флажка и т.д. Эти установки распространяются только на обмен информацией между ГИС и другими программами операционной системы.

#### 14.6. Создание точечных объектов

Выше были рассмотрены процедуры работы с линейными и площадными объектами. На карте существует значительное число так называемых точечных объектов, которые большей частью представляют собой условные знаки.

В ГИС создание точечных объектов на карте осуществляется за счет использования набора библиотек условных знаков, которые поставляются с инструментальными ГИС или могут быть созданы квалифицированным пользователем. Для нанесения условного знака он предварительно активизируется из одной из библиотек, определяется его стиль (цвет, размер и т.п.). Затем устанавливается режим ввода точечного объекта с помощью указателя определяют координатное положение условного знака (точечного объекта).

Для построения новых точечных объектов необходимо модифицировать существующие объекты или включить новую библиотеку точечных объектов.

Одним из простейших способов модификации является использование режима «Стиль символа». Этот режим позволяет (в общем случае) осуществить: выбор начертания, установку размера и цвета, указать угол поворота, выбрать специальные эффекты (тень, подчеркивание, оконтуривание и др.) для символа точечного объекта, изменение атрибутов существующего точечного объекта. ГИС могут поддерживать различные виды символов, например векторные символы, символы из установленных шрифтов TrueType, растровые символы и т.д.

Для создания нового символа можно использовать собственные растровые картинки в качестве Растровых символов. В списке наборов символов они могут находиться под именем «Растровые символы». Для того, чтобы Растровые символы появились в списке «Наборы» диалога «Стиль символа», нужно скопировать растровые картинки в формате Bitmap в библиотеку символов в соответствии с

требованиями этой библиотеки ( размер в байтах, число градаций, цветность и т.д.)

### 14.7. Вывод информации

В режимах вывода информации в ГИС могут возникнуть разные задачи вывода. Традиционной является задача получения карт. Менее известны в технологическом плане другие задачи.

В ходе работ возникает необходимость вывода изображения в виде файла, например для презентации или отчета. Такая возможность реализуется путем экспорта окна. В окне появиться изображение карты, таблицы или того и другого. Эта процедура использует моделирование типа «вектор – растр».

При выполнении экспорта окна выполняют следующие действия:

- определяют размеры изображения экспортируемого из окна,
- задают имя, диска и каталога для нового файла,
- задают формат изображения.

Окончательное представление картографической информации осуществляется с помощью режима «Отчет».

В строке сообщений показывается текущий размер окна Отчета. Размер 37.46% означает, что изображение в окне Отчета составляет 37.46% от реального размера страницы. Размер 123% означает, что изображение в окне Отчета составляет 123% от реального размера страницы.

При создании отчета, если не открыто ни одно окно, создается окно Отчета с чистой страницей. Если же имеются открытые окна, то ГИС выведет на экран диалог.

Открывая новое окно Отчета, ГИС устанавливает размер и ориентацию страницы (обычная или "портрет") на основании текущих параметров для принтера. При каждом последующем открывании окна используются те параметры принтера, которые установлены при открывании окна. Размеры и ориентация объектов в окне при этом не меняются, изменяется только разбиение содержимого окна на страницы.

После создания окно Отчета становится активным окном программы. ГИС добавляет в строку меню пункт Отчет. В меню Отчет входят команды работы с окнами Отчетов.

Затем отчет помещается в рамку, которая корректируется по желанию пользователя путем изменения параметров объекта типа "рамка" (фрейм)



Механизм фреймов позволяет выводить в поле отчета много-оконную информацию: легенды, заголовки, комментарии, фрагменты карт в других масштабах и т.д.

После создания рамки отчета с ней можно работать с ней как с графическим объектом. Пользователь может менять ее размеры и положение, тип линий и штриховки, а также вырезать, копировать и вставлять ее.

Обращаем внимание на то, что рамка отчета может иметь специальные режимы « прозрачности рамки », выравнивания объектов рамки, изменения стиля и т.д. Кроме того, пропорции рамки и тип линии ее создающий можно менять также как и создаваемый отчет

Если внутри рамки помещена карта, то она имеет те же пропорции и масштаб, что и в окне Карты. . Изменение пропорций рамки не влияют на изменение пропорций карты, которую эта рамка содержит.

Для изменения пропорции карты в окне отчета ее надо изменить в окне Карты. В этом случае ГИС автоматически изменит пропорции и в рамке в окне Отчета.

При анализе изображения отчета иногда возникает необходимость изменения его размера. В таком случае для уменьшения изображения Карты, не меняя размеров рамки, используют режимы зуммирования.

При работе с отчетом возможна установка разных режимов его визуализации:

- "Всегда"
- "Только при переходе в окно Отчета"
- "Никогда"

Объекты Отчета (такие как рамки, тексты, линии, прямоугольники и т. п.) имеют свой порядок относительно друг друга от листа макета к пользователю. В соответствии с этим порядком происходит пересечение объектов на листе Отчета (аналогично MS Office ). При этом объект, расположенный выше, закрывает объект, расположенный ниже. Если объекты не пересекаются, то для того чтобы понять, какой из них выше, а какой сдвигают один из них до пересечения со вторым. По тому, какой из них будет перекрывать другой объект, можно определить, какой объект расположен выше.

Стандартно объекты располагаются снизу вверх в порядке создания объектов. Последний создаваемый объект располагается поверх созданных ранее. Для изменения порядка расположения объектов на листе Отчета используют команды «Отчет > Вниз (Down )» и

«Отчет > Наверх( Up)». В первом случае объект опускается под другие, во втором он поднимается наверх и закрывает собой остальные

В Отчете атрибутивные табличные данные представлены в виде таблицы. Таблица показывается с названиями полей вверху. Сразу за ними следуют записи. Количество колонок и строк зависит от размеров рамки, а не от размеров таблицы.

Число колонок и строк, показываемых в рамке, зависит от размера шрифта и ширины колонок таблицы. Чем меньше размер букв, тем больше строк можно вместить в рамку отчета.

При печати информации необходимо выбрать и настроить принтер, а затем распечатать то, что пользователь считает нужным: отчет карту, список, график.

#### **14.8. Методы обработки данных дистанционного зондирования**

В настоящее время данные дистанционного зондирования (ДДЗ) являются самым оперативным источником получения геоинформационных данных. По этой причине они используются в качестве основных источников для поддержания информации ГИС в актуальном состоянии, особенно тогда, когда фактор актуальности играет решающую роль (военная разведка, контроль стихийных бедствий, экологический мониторинг, разведка природных ресурсов и т.д.).

Именно поэтому целесообразно рассмотреть этот вид технологий отдельно. Кроме того, следует констатировать тенденцию взаимного сближения технологий ГИС и обработки данных дистанционного зондирования (ДДЗ) Земли

В технологиях ДДЗ ведущую роль на стадии сбора играют цифровые методы обработки изображений [21, 22].

На рисунке 14.1. показана типовая схема обработки ДДЗ в ГИС-технологиях. Стадии обработки цифровых данных выделены специально, поскольку на этих этапах обработки объем данных, участвующих в обработке, много больше, чем при геоинформационном моделировании.

На первом этапе обработки осуществляется импорт данных, полученных со спутника, либо сканирование данных.

Второй этап предполагает обязательный анализ данных для последующего составления плана обработки изображений.

На третьем этапе осуществляется ректификация изображений. В случае необходимости оно может быть переведено в заданную картографическую проекцию

На четвертом этапе возможно объединение или комбинирование нескольких изображений с целью получения целостной картины об исследуемом объекте или явлении.

На следующем этапе изображение, синтезированное из нескольких других, подвергается обработке по улучшению качества и приведения всех его разных характеристик к единым характеристикам синтезированного изображения.

Для повышения качества дешифрирования объектов и улучшения качества векторизации растрового изображения на шестом этапе осуществляется автоматизированная классификация объектов изображения и их группирование по свойствам ( в дальнейшем атрибутам). Этот этап существенно упрощает процесс организации атрибутивных данных на основе автоматизированного решения этой задачи.

Следующий этап создает основу для применения ГИС-технологий. В нем происходит создание структур атрибутивных данных в соответствии с требованиями конкретной ГИС и создается структура связей позиционных и атрибутивных данных.

На восьмом этапе осуществляется векторизация растрового изображения с использованием данных классификации и организованной связи «координаты-атрибутов». На этом этапе осуществляется существенное ( на 2-3 порядка) сжатие исходных данных при сохранении информативности о выбранных объектах.

На девятом этапе строят цифровую модель как основу хранения данных и моделирования в ГИС.

На десятом этапе осуществляются процедуры геоинформационного моделирования, которые могут повторно включать ряд процедур, таких как комбинирование объектов, ректификация, классификация и др. Цель этих процедур заключается в более углубленном исследовании объектов ГИС.

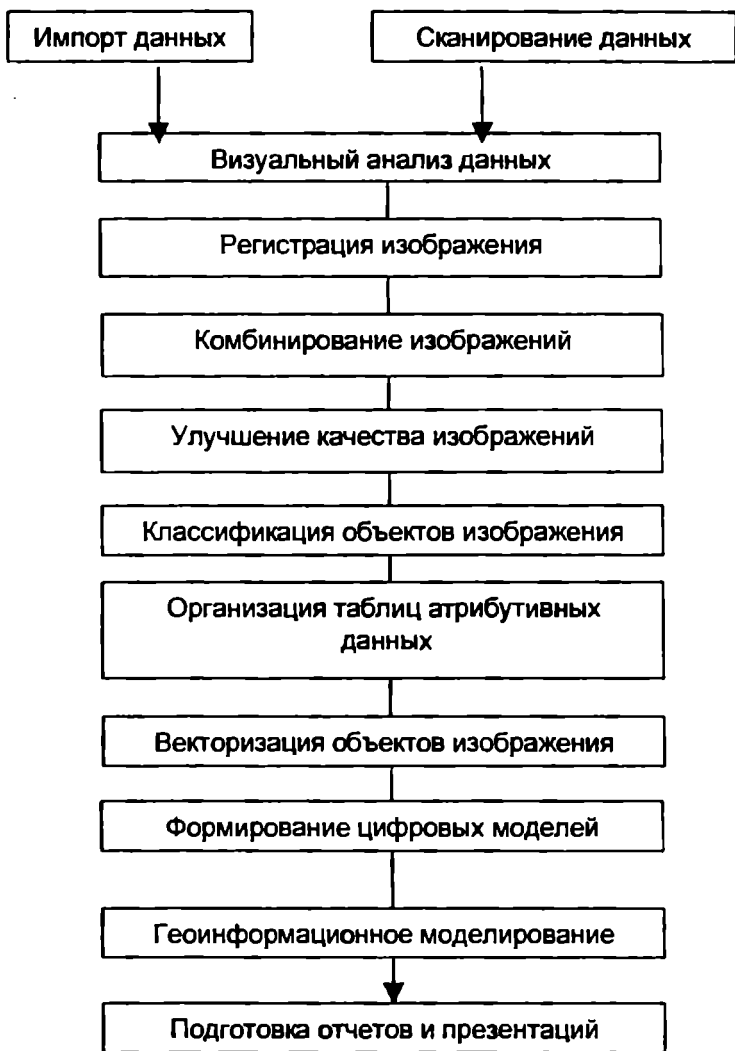


Рис.14.1. Технологическая схема обработки ДДЗ в ГИС-технологиях

На последнем этапе результаты геоинформационного моделирования оформляются в виде отчетов, презентаций, справок, карт и других документов.

Подводя итог рассмотренной технологии, следует обратить внимание на две особенности.

В отличие от *классических* ГИС-технологий, в которых цифровая модель строится по принципу «одна карта – одна цифровая модель», в технологии приведенной на рис. 4.1. возможны другие принципы построения: «несколько разных карт – одна модель», «несколько снимков (разного масштаба) – одна модель», «несколько карт и несколько снимков – одна цифровая модель».

Процессу векторизации предшествует процесс классификации, который не только группирует исходные объекты векторизации, что создает основу для эффективного выбора фильтров векторизации, но и устанавливает структуру атрибутивных данных, также перенося этот процесс (пока только частично) с человека на программу.

В целом уровень автоматизации комбинированной ГИС-технологии, предназначенной для обработки ДДЗ выше, чем многих других ГИС-технологий. Недостатком такой технологии является требование значительных вычислительных ресурсов, мощных программных средств и более квалифицированного пользователя

## 15. ГИС как система управления

ГИС как интегрированная система наследует свойства и возможности других систем. Естественно, что она наследует способности автоматизированной системы управления [8, 46, 93, 98, 105, 106].

В теории управления выделяют три уровня управления: стратегический (высший), тактический (средний), низший (операционный). ГИС как информационная система может быть использована на трех уровнях управления [123].

На *стратегическом* уровне для принятия решений используют информацию, полученную на основе результатов обработки данных в ГИС.

На *среднем* (менеджерском) выбирают варианты и технологии при организации технологий обработки данных.

На *операционном* уровне для принятия решений используют информацию при непосредственной обработке данных оператором с помощью конкретных регламентных методов, входящих в состав геоинформационных технологий.

В настоящее время широко применяются интегрированные информационные системы как эффективное средство реализации новых информационных технологий. К числу таких систем относятся геоинформационные системы (ГИС), которые благодаря своей многофункциональности и многоаспектности продолжают развиваться и находить все более широкое применение во многих сферах деятельности, включая и управление.

По своим технологическим возможностям одним из основных назначений ГИС является управление и принятие решений.

Обобщенная модель управления в геоинформатике показана на рис.15.1.

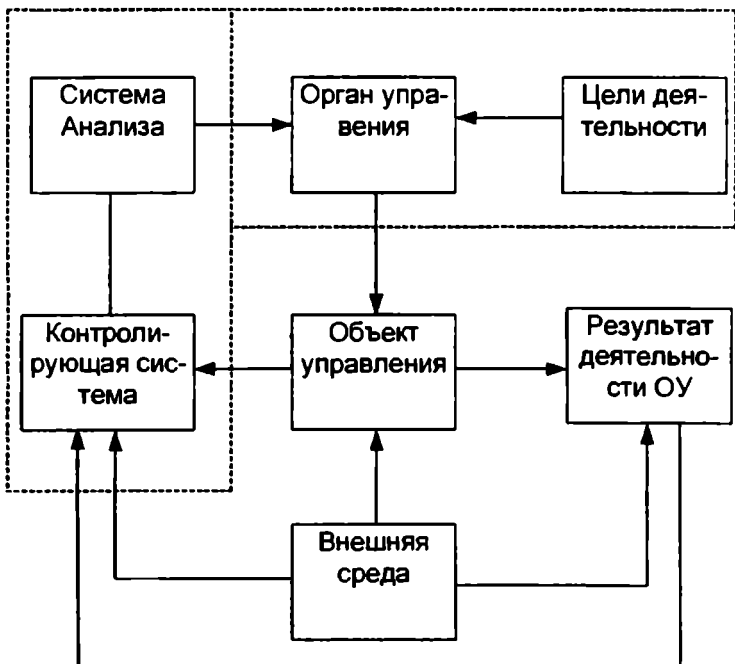


Рис.15.1. Обобщенная модель управления

Особенностью такой модели является наличие цепи обратной связи, включающей не только модель объекта управления, но и модель среды, в которой этот объект находится.

Элементами модели управления являются объект управления (ОУ) и система управления. Объект управления функционирует во внешней среде, которая оказывает воздействие на ОУ и на результат его деятельности (рис. 15.1).

Система управления (СУ) включает контролируемую систему, которая занимается сбором информации о состоянии ОУ, о внешней среде, о результатах деятельности ОУ.

Информация, собранная в контролирующей системе, поступает в систему анализа, где производится обобщение, анализ и представление информации.

Результаты анализа поступают в орган управления, где в соответствии с целями деятельности объекта управления принимается решение и результат принятия решения направляется на ОУ.

Рассматривая данную систему по трем уровням управления можно определить, что на нижнем (операционном) уровне осуществляется сбор первичной информации и ее унификация (контролирующий орган рис. 15.1).

На среднем уровне управления осуществляется обобщение, анализ, получение оценок и т.п. Эти работы выполняются в системе анализа (рис. 15.1).

На верхнем уровне управления (орган управления) принимаются решения.

Таким образом, в процессе управления задействованы все уровни управления: нижний, средний и высший.

Модель поддержки принятия решений представляет собой фрагмент модели принятия решений. В процессе поддержки принятия решений обработка информации выполняется в основном на среднем и (в меньшей степени) на операционном уровнях.

Принципиальным отличием технологии поддержки принятия решений от технологий управления является подготовка информации (моделей), на основе которых впоследствии будет приниматься решение.





Рис. 15.2. Схема поддержки принятия решений

На рис. 15.2 приведена технологическая схема поддержки принятия решений, основанная на применении аналитической информационной системы.

Аналитическая информационная система (АНИС) представляет собой объединение контролирующего органа и системы анализа (рис. 15.1)

Это объединение не механическое, а технологическое, т.е. обусловлено целями и технологиями обработки информации.

АНИС исследует внешнюю среду и выделяют два класса факторов:

- независимые факторы (НФ), на которые невозможно воздействовать в рамках системы управления;
- изменяемые факторы (ИФ), которые можно изменить применяя управляющие решения.

АНИС должна содержать базу данных, включающую в себя набор типовых методов принятия решений для так называемых регламентных (известных) ситуаций.

По мере анализа новых ситуаций и получения для них эффективных решений они переводятся в класс регламентных. Наличие независимых факторов, а также неопределенности в состоянии среды исключает возможность однозначного принятия решения.

Уменьшение неопределенности осуществляют разными способами, одним из которых является моделирование и получение наборов условных состояний внешней среды. Среди этих возможных состояний выделяют существенные (значимые) и несущественные. Для каждого из значимых состояний среды разрабатывают набор проектов решений.

В этом и заключается поддержка принятия решений.

Таким образом, поддержка принятия решений может быть определена как совокупность технологий:

- сбора, обобщения, анализа;
- выявления изменяемых и неизменяемых факторов внешней среды;
- оценка неопределенности ситуации внешней среды;
- уменьшение неопределенности путем моделирования состояний внешней среды;
- использование базы типовых решений для регламентных ситуаций;
- разработка набора альтернатив решений;
- разработка прогнозных оценок.

Следует подчеркнуть наиболее важные моменты: уменьшение

неопределенности путем моделирования (создания набора) ситуаций и разработка прогнозных оценок.

Рассматривая ГИС как аналитическую систему для поддержки принятия решений, следует отметить, что она предоставляет дополнительную возможность лицу принимающему решение (ЛПР).

В отличие от набора альтернатив  $A_1, A_2, A_3, \dots$  (рис. 15.2), которые можно получить в существующих АИС, ГИС дополнительно обеспечивает набор визуальных средств поддержки принятия решений  $B_1, B_2, B_3$ , (рис. 15.2). К числу таких средств относятся тематические карты и планы, содержащие не только изобразительные характеристики, но и метрические. Метрические характеристики служат основой для точных расчетов и соответственно для экономических оценок того или иного решения.

Можно констатировать, что поддержка принятия решений (как технология), применяется при наличии неопределенности в исходной информации [102, 111, 112].

Технология поддержки принятия решений сводится к дифференциации состояний внешней среды и объекта управления. Неопределенность уменьшается или исключается путем ввода набора параметров, характеризующих набор состояний, и определения этих параметров на интервальных (от и до) или точечных (ровно) значениях. Набор возможных состояний среды и объекта определяет набор альтернатив [102, 106, 111, 123].

После разработки набора альтернатив может возникнуть задача уточнения или определения параметров состояний среды. Это определится на основе сопоставительного анализа альтернатив.

В этом случае проводят дополнительные работы для получения достоверной информации, т.е. переход от смоделированной ситуации к реальной. Данная задача, в свою очередь, связана с определением стоимости достоверной информации, которая может быть получена в результате исследований.

Стоимость информации определяется по различным критериям, отметим три из них:

- трудозатратный;
- получения заданной прибыли;
- уменьшения возможных убытков.

Следует отметить, что первый критерий наиболее простой, так как он определяет стоимость информации на основе затрат на ее создание или получение. Затраты определяются стоимостью работ соответствующих специалистов.

Рассмотрим три критерия на примере.

Некая фирма собирается создать систему оптимального размещения товаров. Планирование этого процесса можно выполнять вручную (на основе опыта работающих специалистов) или автоматизированно (на основе использования программного обеспечения). Такая ситуация соответствует двум альтернативам: А1 - ручная обработка, А2 - автоматизированная обработка.

Для оценки стоимости программы в альтернативе А2 можно использовать трудозатратный критерий. Он позволяет определить стоимость информации (в нашем случае программы) равной оплате труда программиста по ее созданию. Этот подход дает первую оценку стоимости -  $C_1$ . Программа в процессе работы может приносить ежемесячную прибыль в размере  $\Pi_m$ .

Как показывает опыт жизненный цикл программной продукции не превышает 3 лет. Общая прибыль  $\Pi_0$ , приносимая программой за период ее функционирования, определяется как

$$\Pi_0 = 36 \cdot \Pi_m - 3,$$

здесь 3 - затраты на получение или создание программы.

Второй критерий определяет стоимость информации по величине максимальной прибыли

$$C_n \leq \Pi_0$$

за период эксплуатации средства, использующего данную информацию.

Программа может быть направлена не только на получение прибыли, но и на выявление и предотвращение нежелательных рискованных или неблагоприятных ситуаций [106]. В этом случае используется третий критерий - оценка стоимости информации. Если программа выявляет и предотвращает возможные убытки так, что в среднем месячная величина предотвращенных убытков составляет  $Y_m$ , то за период жизненного цикла (при условии относительного постоянства внешней среды) общая сумма предотвращенных убытков  $Y_0$  составит

$$Y_0 = 36 \cdot Y_m - 3,$$

3 - затраты на покупку или разработку информационного продукта. Третий критерий определяет стоимость информации как

$$C_y \leq Y_0.$$

Вариации с этими критериями при учете конкретных условий работы позволяют оценить стоимость информации и сопоставить их с затратами на ее получение.

Для поддержки принятия решений, включая анализ стоимости

информации, эффективно использование ГИС.

ГИС как интегрированная система включает возможности многих систем, таких как АСУ и др [8, 98]. Эти возможности могут быть использованы только при создании соответствующей технологии, т.к. исходный пакет ГИС, как правило, представляет собой инструментальную систему, обеспечивающую общие методы анализа и интерпретации данных.

По этой причине следует подчеркнуть, что при решении картографических задач современные инструментальные ГИС позволяют получать их решение полностью. Однако при переходе к задачам управления ГИС представляет собой лишь ядро информационной системы, которое требует создания дополнительной интеллектуальной оболочки [104].

Из сказанного следует, что ГИС может быть системой поддержки принятия решений, но для повышения ее эффективности при управлении интерфейс ГИС должен быть дополнен средствами анализа, применяемыми, например, в маркетинговых и других системах.

## 16. Земельный кадастр

### 16.1. Кадастр как система налогообложения

Термин "кадастр" часто ошибочно отождествляют с термином "земельный кадастр". В общем случае кадастр это реестр, содержащий перечень об объектах налогообложения, к числу которых принадлежит также земля.

В нашей стране применяют три основных вида налогов:  
 кадастровый,  
 на основе декларации,  
 "у источника".

В первом случае объект налогообложения классифицирован по определенному признаку и разбит на группы. Перечень этих групп и их признаки заносится в специальные справочники. Для каждой группы установлена индивидуальная ставка налога. Такой подход характерен тем, что величина налога не зависит от доходности объекта.

Во втором случае налогоплательщик заполняет декларацию, в которой он приводит расчет дохода и налога с него. Характерной чертой такого метода является то, что выплата налога производится после получения дохода и лицом получающим доход.

В третьем случае налог вносится лицом, выплачивающим доход (бухгалтерия). При этом методе оплата налога производится до получения дохода, причем получатель дохода получает его уменьшенным на сумму налога.

Таким образом, земельный кадастр одной из своих основных функций имеет сбор налога с земельных объектов.

Понятие кадастра включает систему метрических и юридических данных. Эти данные необходимы для налогообложения земельных участков, собственности, а также регистрации юридических прав [1, 14, 34, 52, 69, 121].

### 16.2. Исторический обзор развития кадастра.

Древний Египет. Исторические корни возникновения земельного кадастра уходят в глубокую древность. Элементы земельного кадастра можно проследить в ранних земледельческих поселениях вдоль рек Тигр, Евфрат и Нил, где доходы для Фараона и жрецов собирались главным образом от налогообложения земельного дохода. Для получения такого налога необходимо измерение земельных участков, т.е. проведение кадастровой съемки (3000г до н.э.). Резуль-

таты съемок отображались в то время на глиняных таблицах, с указанием границ земельных участков и их площадей.

**Греция и Римская империя.** Позже греки и римляне создали более сложные системы земельных данных и съемок для налогообложения земли. Земельный кадастр Древнего Рима представлял собой описание земельной собственности, начало которого приписывается Сервию Туллию (VI в. до н.э.). Сведения о размере земельных участков, способе их обработке, качестве и доходности земель вносили в специальные реестры. На бронзовые таблички наносили планы имений, их названия, границы и размеры землевладения.

Появление термин а "кадастр" относят ко времени Римского правителя Августа (27 год до н.э. - 14 год н.э.). При нем была утверждена единица учета сбора дани за землю, названная "capitigum" и введена перепись населения "capitum registrum". Со временем эти слова слились в одно- "capitastrum" и затем в "catastrum".

Первый земельный кадастр Римской империи включал работы по переписи населения, разделению земель на участки (центурии), точному измерению их границ, определению качества земель, составлению карт этих участков и нанесению на камни отображения их границ, номеров, юридического статуса, имени владельца и суммы налога представляли собой.

Создание и ведение кадастра строго регламентировалось инструкциями и указаниями. Эти инструкции регламентировали проведение съемок и изменение юридической терминологии. Как видим, уже в то время существовала правовая поддержка ведения кадастра.

Следует сделать вывод, что в рассмотренные периоды кадастр применялся как инструмент внутренней политики. Главную роль в этом кадастре играет **имущественный фактор**, определяющий владельца и его обязанности как налогоплательщика внутри государства.

**Средние века.** В средние века содержание кадастровых съемок на основные принципы ведения земельного кадастра не влияло. Съёмка земель оставалась чисто линейной и проводилась с помощью веревки и жезла, слабо использовались геометрические знания и вычисления. Подобные методы использовались вплоть до XVIII века.

В средние века в период межгосударственных войн значение кадастра выходило за рамки одного государства. Главную роль начинает играть **юридический фактор**, который определяет правовой аспект владения территорией.

Границы государства отмечались пограничными столбами, которые назывались марками. Граф, в чьем ведении находились по-

границные земли империи (государства), наделялся особыми правами и получал титул **маркграфа**.

Для установления имущественных и юридических прав на земельные владения применялись процедуры:

законодательное установление прав на земельный участок;

демаркация (от слова марка) - установление грании на местности;

съемка местности;

регистрация владения в специальные регистрационные книги.

Начиная с 1718 г. Джованни Джакомо Мариони (Милан) разработал первый кадастр на основе строго соблюдения научных методов, с определением границ участков методом триангуляции и полигонометрических сетей. Этот кадастр содержал в качестве документов ситуационные карты всех сельских общин, выполненные с использованием мензулы в масштабе 1:2000. На картах отображались **парцеллы** (участки) отдельных общинных владений. При этом определялись площади владений, типы грунтов, чистая прибыль, которая являлась основой для обложения налогом.

Развитие данной кадастровой системы можно проследить от налогового картирования итальянских провинций Милан и Мантуа. Благодаря своему качеству и точности, этот кадастр, называемый Миланским, был введен в действие 1 января 1760 года и служил образцом для разработки в XIX веке кадастров Франции, Бельгии, Австрии, Голландии и Швейцарии.

В Австрии выполнили кадастровую съемку с 1785 по 1789гг. по всей территории страны, включая Австро-Венгрию.

В 1807 году Наполеон назначил математика Деламбера председателем кадастровой комиссии. Задачей комиссии была съемка свыше ста миллионов участков, классификация участков по плодородию почв и оценка продуктивной возможности участков. Помимо этого проводили следующие процедуры:

- опись различных участков с именами владельцев собственности;
- определение основ их общей продуктивной возможности;
- оценка общего дохода;
- проведение налогообложения данных, которые должны служить основой для будущего налогообложения.

В настоящее время в большинстве стран Европы система регистрации земель также базируется на принципе наделов (parcel-based).



### 16.3. Развитие Кадастра в России

Первые сведения о кадастре в России относятся к X веку и связаны со сбором поземельного налога и оценкой земель. Первое упоминание о ведении работ по межевому (кадастровому) картографированию для отвода земель Святогорскому монастырю на Псковщине относится к 1483 году.

В русской картографии XV-XVII веков кадастровое картографирование было выделено как самостоятельное направление и называлось **межевым**. В описаниях земель, собранных в постовых, смотренных, дозорных и межевых - книгах содержался и картографический материал [13]. Кроме того, эти материалы дополнялись данными результатов натурных землемерных работ.

**Землемерные работы** заключались в измерении длин граничных линий "мерной вервью" (специальной веревкой длиной в 80, 40 и 30 сажений), которые разделяли землю по категориям угодьям, пашням, лугам, лесам и по качеству (добротности) на "добрую", "среднюю" и "худую".

Земельные планы, создаваемые при межевании, являлись юридическим документом на закрепление права владельцев на землю.

В начале XIX века основные работы по межеванию в России были завершены. В дальнейшем проводились работы по проверке и обновлению данных межеваний.

Еще в XVIII веке в городах России проводились **инвентаризационные работы**. Вышедший в 1768 году Указ гласил "о создании специальных планов по всем городам, по их строению и улицам в каждой губернии"[1, 13].

В книге чертежей и рисунок опубликовывались планы перепланировки городов России. Во многих городах, до установления советской власти, проводились оценка недвижимого имущества и его опись. Кроме того, составлялась схема земельного участка на каждое домовладение.

Итак, обобщим систему земельного кадастра в дореволюционной России. Исходя из вышеизложенного можно утверждать, что главный фактор российского кадастра имущественный.

При этом система земельного кадастра включала следующие составные:

- развитие специального направления картографирования, которое называлось **межевым**;
- кадастровые съемки, называемые **землемерными работами**. Эти съемки дополнялись проведением специальных расчетов и вычислений;

- систему классификации земель по территориальным и качественным признакам;
- систему обновления и инвентаризации недвижимости и земельного имущества;
- создание специальных кадастровых планов.

В настоящее время кадастр также поддерживается указами президента, постановлениями правительства и специальных ведомств. Эти документы составляют правовую основу ведения земельного кадастра в нашей стране.

#### **16.4. Современное состояние кадастра за рубежом**

В настоящее время кадастр ведется во всех странах мира. Он постоянно совершенствуется и неразрывно связан с понятием учета, оценки, состояния и использования различных природных ресурсов, инженерной деятельности, экологии. Кадастр характеризуется выделением внутренне однородных по своим условиям территориальных единиц, картографированием и составлением описания количественных и качественных характеристик этих единиц. В зарубежной практике понятие "кадастр" довольно часто связывают с понятием "земельный реестр".

В настоящее время в Западной Европе существует два основных направления ведения кадастра:

- юридический кадастр (Германия);
  - фискальный (налоговый) кадастр (Франция).
- Рассмотрим подробно эти два вида кадастра.

Юридический кадастр имеет в качестве основных функций оформление земельных прав и управление использованием земли. В него включаются четыре основных операции, каждая из которых обеспечивает земельную информацию:

- **решение** - законодательное установление прав на земельный участок;
- **демаркация** - установление границы каждого участка на местности;
- **съёмка** – местности;
- **регистрация** - ввод данных в регистрирующую систему, например в регистрационные книги

Эти операции взаимосвязаны, изменение в одной части может вызвать изменение в других частях.

Хотя местоположение границ играет важную роль в юридическом кадастре, требования к точности в этом кадастре невысокие.

Основной целью его является система регистрации прав на земельные участки.

Юридический кадастр имеет две основные части: атрибутивную и метрическую. В первой части делается опись или реестр, состоящий из информации о каждом участке, такой как имя владельца и права собственности на землю. Вторая часть имеет ссылку на первую и содержит детальное описание участка в виде карты или схемы измерений.

Основными функциями юридического кадастра являются:

- предоставление достоверных данных о собственности;
- сокращение земельных споров;
- совершенствование передачи прав собственности;
- стимулирование земельного рынка;
- управление государственными землями.

Фискальный кадастр можно определить как систему инвентаризации земельных участков, необходимую для определения стоимости каждого участка и величины налогового сбора. Этот кадастр имеет следующие основные функции:

- информационной основы для налогообложения;
- поддержки программ финансового распределения;
- поддержки земельного рынка;
- развитию управления землепользованием;
- обеспечения земельной информацией.

## **16.5. Государственный земельный кадастр**

Существует несколько видов кадастра: земельный, городской, водный и др. Земельный кадастр направлен на решение задач исследования и правильного использования земельных ресурсов. Он имеет следующие основные функции:

- обеспечение рационального использования земельных ресурсов страны;
- улучшение сельскохозяйственного производства;
- охрана земель от эрозии, заболачивания, засоления и других вредных процессов, снижающих плодородие почв;
- решение задач об изъятии земель для не сельскохозяйственных нужд.

*Государственный земельный кадастр* - единая, государственная система признания и удостоверения государством факта возникновения, существования или прекращения существования объектов кадастрового учета. Внесенные в учетные формы сведения об объекте кадастрового учета являются единственным доказательством су-

уществования учтенного объекта в границах, признанных при его учете.

Основной целью создания государственного земельного кадастра является формирование, накопление и обновление сведений о земле в интересах реализации, поддержания и регулирования установленных в государстве земельных отношений и правоотношений.

Действующее законодательство определяет, что земельный кадастр ведется Государственным земельным комитетом и его территориальными органами. В соответствии с этим кадастр имеет иерархическую структуру. Перечислим уровни этой иерархии:

федеральный государственный земельный кадастр,

государственные земельные кадастры субъектов Российской Федерации,

государственные земельные кадастры муниципальных образований.

В процессе государственного кадастрового учета объекту присваивается кадастровый номер. Государственный кадастровый учет осуществляется по месту нахождения объекта кадастрового учета в соответствующем земельном органе. Учетными единицами в государственном земельном кадастре являются земельные участки и территориальные зоны.

Кадастровые номера служат основой государственного кадастра. Административно-территориальное деление Российской Федерации производится на 89 Субъектов Федерации (21 республика, 6 краёв и 49 областей, 2 города - Москва и Санкт-Петербург, 10 автономных округов и 1 автономная область). Номер (код) Субъекта Федерации (№№ 1-89) составляет первую позицию Кадастрового Номера (КН) любого земельного участка или иного объекта недвижимости на территории России.

Характерным примером применения такого подхода является то, что вновь выдаваемые номера на автотранспортные средства включают кадастровый номер субъекта федерации, в пределах которого регистрируется данное транспортное средство. Исключение составляет Москва, где наряду с кадастровым номером 77 выдаются номерные знаки с номером 99.

## **17. Земельные информационные системы**

Исследование и использование природных ресурсов, рациональное ведение народного хозяйства, охрана природы и мониторинг, принятие важных практических решений, связанных с окружающей средой, невозможны без информационных систем и надеж-

ного информационного обеспечения. Создание земельных информационных систем (ЗИС) служит основой современных методов землепользования [121].

Применение ЗИС повышает эффективность исследования и использования природных ресурсов, рационального ведения хозяйства, охраны окружающей среды.

ГИС являются идеальным инструментом для создания земельных информационных систем. По существу современные ЗИС создаются только на основе инструментальных пакетов геоинформационных систем и на основе геоинформационных технологий.

Таким образом, современная ЗИС является многофункциональной и направлена на решение задач землепользования. Технологически эта система представляет собой специализированную ГИС, ориентированную на выполнение функций ЗИС.

Однако для реализации ЗИС на основе геоинформационных технологий необходимо предусмотреть создание системы правовой поддержки принятия решений, что в большинстве инструментальных пакетов ГИС отсутствует. Эта система может быть встроена в ГИС или связана с ней через интерфейс работы с удаленными базами данных.

Земельную информационную систему можно рассматривать как подсистему системы для принятия решений в отношении землепользования и землевладения в стране. Одновременно она является подсистемой социально-экономической системы. Поэтому одна из основных функций ЗИС - управление на различных уровнях.

В настоящее время выделяют три основных уровня управления: стратегический (высшее руководство), тактический (менеджеры), операционный (рядовые исполнители) [90, 116, 121, 123].

Особенность ЗИС заключается в возможности ее применения на любом из уровней управления. Однако при этом ее функции и методология существенно отличаются.

### **17.1. Применение ЗИС на операционном уровне управления**

Нижний уровень - это операционный, или эксплуатационный. На этом уровне решаются регламентные задачи. Выдаются справки и типовые документы в соответствии с законодательством.

Рядовые потребители заинтересованы в том, чтобы получить оперативно необходимую информацию в форме и в виде, соответствующем их запросам.

На данном уровне главная функция общей земельной информационной системы состоит в генерировании и распределении ин-

формации, связанной с землепользованием, между ее потребителями. Для того, чтобы реализовать эту общую функцию, нужно осуществить ряд функций более низкого порядка: данные должны быть собраны, введены в систему, сохранены и обработаны. Эти действия исполняются независимо от того, является система ручной или автоматической.

Повышение эффективности ЗИС на этом уровне может быть достигнуто в процессе сбора пространственных данных. К ним, например, относятся более точные методы наземной съемки, использование GPS, фотограмметрии и дистанционного зондирования.

Эффективность ЗИС при сборе данных не обеспечивается не только техническими средствами. Рост эффективности возможен также за счет организационных изменений.

Например, кадастровая съемка может выявить, что учет земель и землевладения находится под разными “перекрывающимися” юрисдикциями [19]. Разрешая эти противоречия и упрощая правовой процесс, касающийся учета земель, можно значительно повысить эффективность и экономичность земельной информационной системы.

Одно из наиболее широко известных преимуществ, связанных с отдачей ЗИС, состоит, например, в том, что она позволяет уменьшить избыточность информации за счет центрального агентства, ответственного за обработку исходных карт вместо нескольких агентств, занимающихся классифицированием одной и той же территории в разных масштабах и с разной точностью.

Особо следует отметить прирост эффективности и экономичности за счет стандартизации. Это касается и земельной информации и способов ее классификации и способов кодирования и хранения данных.

Основной функцией ЗИС операционного уровня при решении задач землевладения является обеспечение информацией о правах на пользование землей и способность увязать облагаемую налогом собственность с владельцами и с ограничениями, налагаемыми на эту собственность.

Это означает, что система должна установить облагаемый налогом статус владельцев каждой части пакета прав, а также права, которые обеспечены правилами, распространяющимися на эту собственность. Кроме того, система позволяет обрабатывать земельные сделки, обеспечивать права и отвечать на вопросы, касающиеся владения, сервитутов, дополнительных прав и общественных интересов [19].

Функциями ЗИС операционного уровня, при решении задач землепользования, является способность системы выделять различные характеристики земли. Например, тип почвы, растительный покров, близость к чему-либо и т.д., а также способность увязывать эти характеристики между собой и с нормами охраны окружающей среды, в особенности применительно к поймам, заливаемым лугам, основным сельскохозяйственным землям и т.д.

### **17.2. Применение ЗИС на тактическом уровне управления**

Второй уровень управления - тактический. На этом уровне осуществляется анализ и разрабатываются новые регламентные методы работы, которые затем утверждаются на высшем уровне. На нем же осуществляется подготовка предложений по изменению тактики работы. На этом уровне одной из главных задач является подготовка принятия решения.

При этом подготавливаемые решения должны иметь юридическую силу. Таким образом, при разработке ЗИС для среднего уровня управления они должны быть обеспечены юридическими нормами принятия решений или подсистемами правовой поддержки решений.

Второй уровень обеспечивает анализ и отправку информации либо на низший, либо на высший уровни. Кроме решения типовых задач по обработке информации на нем осуществляется обобщение данных и выявление тенденций.

### **17.3. Применение ЗИС на стратегическом уровне управления**

Высший уровень управления - это стратегическое управление. Информация поступает на данный уровень в обобщенно-аналитической форме.

Руководство, использующее земельные информационные системы, нуждается в большей осведомленности и заинтересовано в повышении уровня принятия решений своих общественных служб или граждан. Поэтому на этом уровне ЗИС должны включать в свой состав или иметь выход на системы анализа статистической информации (пакет "Статистика") или на оперативно-аналитические системы типа OLAP [123].

Применение ЗИС на третьем уровне управления дает возможность воздействовать на способ землепользования и землевладения в стране. Поэтому при разработке земельных информационных систем необходимо, чтобы они были совместимы с другими инструментами правительственной политики, направленной на улучшение исполь-

зования и владения землей. И наоборот, следует стремиться к тому, чтобы другие инструменты функционировали совместно с земельной информационной системой.

Следовательно, при разработке земельной информационной системы нужно уяснить, как она стыкуется с более широкой системой землевладения и обмена землями.

При разработке ЗИС для высшего уровня управления необходимо рассматривать ее как один из многих инструментов государственной политики воздействия на использование и распределение земли.

По этой причине ЗИС должна быть образом интегрирована и совместима с другими правительственными системами и программами, например программами налогообложения частной собственности, перераспределения, регулирования и сокращения степени вмешательства государства или уплотнения земель.

В противном случае, когда проект ЗИС не интегрирован с другими правительственными программами, предполагаемая программа ее использования может работать в прямо противоположном направлении. Именно поэтому при разработке земельных информационных систем важно рассматривать их в контексте с более широкими стратегиями и обеспечивать их взаимную интеграцию и совместимость.

С этой целью при разработке и применении земельной информационной системы приходится вносить изменения, обеспечивающие ее совместимость с такими программами, а иногда менять сами программы для того, чтобы приспособить их к новой ситуации, возникающей в связи с применением ЗИС.

## **18. Геоинформационное прогнозирование**

Проблема прогнозирования, из-за большого количества неуправляемых и слабо предсказуемых внешних факторов является одной из наиболее сложных в геоинформатике.

Комплексное исследование окружающей среды в настоящее время эффективно на основе применения ГИС-технологий и геоинформационных систем. В основу исследования заложен анализ природной и социальной среды как совокупности геоинформационных объектов. Такой подход приводит к появлению новой технологии прогнозирования - геоинформационного прогнозирования [100, 106].

Геоинформационное прогнозирование представляет собой набор методов получения и анализа информации об окружающей среде и разработки прогнозных оценок для поддержки принятия ре-



шений. Целью геоинформационного прогнозирования является снижение уровня неопределенности при принятии решений.

Геоинформационное прогнозирование решает две основные задачи: оценку тех или иных геоинформационных параметров для данного момента времени и получение прогнозных оценок на перспективу.

Оценка текущих параметров применяется для оперативного анализа существующей ситуации. Прогнозные оценки используют при оценке альтернатив принимаемых решений или при изучении явлений и их будущих последствий.

Кроме того, геоинформационное прогнозирование как технология интегрирует статистические методы прогнозирования, методы деловой графики, методы геоинформационного моделирования [99] и цифрового моделирования [20, 54, 71, 74, 98, 109, 130].

При геоинформационном прогнозировании можно выделить три качественных этапа обработки информации:

1. Сбор, группировка, обобщение и унификация первичных данных;
2. Анализ, моделирование вторичных (унифицированных) данных;
3. Получение прогнозных оценок и их верификация.

### **18.1. Анализ, моделирование вторичных (унифицированных) данных**

На этом этапе при анализе осуществляется: выбор объектов прогноза; исследование фона (среды); классификация событий; формирование задачи и генеральной цели прогноза.

Анализ унифицированных данных включает предварительное обобщение и группировку данных и построение моделей объекта прогнозирования.

На основе этих данных, после их унификации, создают модели объектов прогнозирования, которые определяются совокупностью цифровых моделей: цифровые модели местности (ЦММ), цифровые модели объектов (ЦМО), цифровые модели явлений (ЦМЯ) [98].

Особенность решения задач прогнозирования заключается в необходимости исследования взаимодействия объекта прогнозирования с внешней средой. Это требует дополнительного построения модели внешней среды [106, 116].

При наличии моделей объекта прогнозирования и данных об окружающей его среде проводится дальнейший анализ и обобщение. На этой стадии необходимо, исходя из задачи прогнозирования, выделить и определить следующие понятия: объект прогнозирования,

значащие переменные прогнозирования, эндогенные и экзогенные переменные.

**Объект прогнозирования** - совокупность качественных и количественных признаков, отражающих свойства объекта прогнозирования в контексте задачи прогнозирования.

**Значащие переменные объекта прогнозирования** - это переменные или показатели в отличие от прочих оказывающие существенное влияние на результаты прогноза.

**Эндогенные переменные** объекта прогнозирования - это значащие переменные, отражающие, главным образом, собственные свойства объекта прогнозирования.

**Экзогенные переменные** объекта прогнозирования - это такие значащие переменные, которые отражают собственные свойства среды ( прогнозного фона).

## 18.2. Методы прогнозирования

Методы прогнозирования можно классифицировать на *эвристические* (при применении которых преобладают субъективные начала) и *экономико-математические* (при применении которых преобладают объективные начала), к числу которых относятся статистические методы.

**Эвристические методы прогнозирования** предполагают, что подходы, используемые для формирования прогноза, не изложены в явной форме и неотделимы от лица, делающего прогноз. При разработке эвристических методов доминируют интуиция, опыт, творчество и воображение. К данной категории методов относятся методы социологических исследований, экспертные методы и методы прогнозирования по аналогам (метод прецедентов). Особенностью этих методов является их зависимость от лица, делающего прогноз.

**Логико-математические методы прогнозирования** основаны на построении формализованных логико-математических моделей. Это позволяет осуществлять обработку данных и получение прогноза на основе алгоритмов без участия субъекта. Особенностью этих методов является их независимость от лица, делающего прогноз. Они могут быть воспроизведены другими лицами, которые неизбежно приведут к получению такого же или близкого прогноза.

К данной группе относятся методы прогнозирования по аналогам (метод моделей), функционально-логическое прогнозирование, структурное прогнозирование, параметрическое прогнозирование, комплексное прогнозирование.

Можно выделить два способа разработки прогнозов, основанных на методах математической статистики: экстраполяцию и мол-

лирование.

В первом случае в качестве базы прогнозирования используется прошлый опыт, который пролонгируется на будущее. Во втором случае строится прогнозная модель, характеризующая зависимость изучаемого параметра от ряда факторов, на него влияющих. Она связывает условия, которые, как ожидается, будут иметь место, и характер их влияния на изучаемый параметр.

Каждый из методов прогнозирования обладает определенными достоинствами и недостатками. Их применение более эффективно в краткосрочном прогнозировании. Многие методы сильно упрощают реальные процессы, что затрудняет возможность отражения в моделях долгосрочного прогнозирования структурных сдвигов и других долговременных тенденций. Для повышения качества прогнозирования необходимо учитывать сезонные, случайные, оперативные и долговременные факторы.

Следует подчеркнуть, что качественному методу также присущи значительные погрешности, поэтому интуиция должна проверяться с помощью доступных фактов и знаний. Следовательно, лишь совместное использование этих двух подходов может обеспечить максимально точный прогноз.

### **18.3. Прогнозные модели и их характеристики**

При проведении прогнозирования может создаваться **прогнозная модель** - модель, исследование и использование которой позволяет получить информацию о возможных состояниях объекта в будущем и (или) путях и сроках осуществления этих состояний.

Геоинформационные данные, как известно, носят временной характер. В том случае, когда прогноз необходим на определенный период времени в будущем, в качестве исходных данных используют временные ряды или временные модели.

В этих случаях прогноз основывается на ретроспективном анализе данных и определении ряда характеристик. К их числу относится **прогнозная ретроспектива**.

**Прогнозная ретроспектива** - процедура моделирования "назад" на некий временной период прогнозирования, на котором исследуется объект прогнозирования и прогнозный фон (среда) с целью получения их систематизированного описания.

**Период упреждения прогноза** - промежуток времени, на который разрабатывается прогноз.

**Период основания прогноза** - промежуток времени, на базе которого строится ретроспектива.

**Прогнозный горизонт** - максимально возможный период

оценки прогноза заданной точности и достоверности.

#### **18.4. Верификация прогноза**

Получение прогнозных оценок должно быть подвергнуто верификации для определения их надежности. Цель верификации прогноза заключается в оценке его функциональной полноты, точности и достоверности. Применяют разные виды верификации.

**Прямая верификация прогноза** - верификация путем разработки того же прогноза другим методом.

**Косвенная верификация прогноза** - верификация путем сопоставления его с прогнозом или данными полученными из других источников.

**Инверсная верификация** - верификация прогноза путем проверки адекватности прогностической модели в ретроспективном периоде.

**Консеквентная верификация** - верификация путем аналитического или логического выведения прогноза из ранее полученных прогнозов.

**Верификация оппонентом** - верификация путем опровержения критических замечаний оппонента по прогнозу.

**Верификация экспертом** - верификация сравнением прогноза с мнением эксперта.

После выполнения и верификации прогноза необходимо оценить его качество. Для этой цели используют следующие показатели:

**Полнота прогноза** - доля вариантов прогноза из множества возможных и /или доля функций объекта прогнозирования, рассмотренных в процессе прогнозирования.

**Точность прогноза** - оценка доверительного интервала прогноза для заданной доверительной вероятности.

**Достоверность прогноза** - оценка доверительной вероятности осуществления прогноза для заданной точности (доверительного интервала).

**Ошибки прогноза** - апостериорная величина отклонения прогноза от действительного состояния объекта.

**Источник ошибки прогноза** - фактор, могущий привести к появлению ошибки прогноза.

#### **18.5. Выбор метода прогнозирования**

При выборе метода прогнозирования следует учесть для какого метода управления предполагается использовать прогноз. Это объясняется тем, что различные типы управления предъявляют разные

требования к виду результатов и точности прогнозирования.

**Текущее управление.** При таком управлении предполагается, что последствия управляющего воздействия будут аналогичны ранее наблюдавшимся при управлении другими объектами. В таких случаях используют прогнозирование по аналогии (прецедент или модель).

**Системное управление** предполагает необходимость анализа множества элементов объекта и среды, а также связей между ними. При таком типе управления чаще всего используют экспертное, функционально-логическое, структурное прогнозирование.

**Ситуационное управление** предполагает необходимость в прогнозе последствий принимаемых решений. Результат такого прогноза может носить качественный (хуже, лучше или предпочтительнее, недопустимо и т.д.) или количественный характер. Поэтому такой тип управления чаще должен использовать экспертное, функционально-логическое, структурное или математическое прогнозирование.

## 19. Геомаркетинг

Необходимость расширения и повышения эффективности маркетинговых технологий привела к появлению геомаркетинга.

Геомаркетинг возник как синтез технологий поддержки принятия решений, технологий визуальной обработки информации, маркетинговых технологий, интеграции различных данных и их комплексной обработки.

Сами по себе геоинформационные системы являются системами визуальной обработки информации и глобальной интеграции данных [106].

Соотношение между маркетингом и геомаркетингом на трех различных уровнях : концепций, методологии и технологии - показано на рис. 19.1. Геомаркетинг включает те же концепции, что и обычный маркетинг. Его возникновение стало возможным благодаря появлению новых информационных технологий и информационного маркетинга. Методология геомаркетинга основана на методологии информационного маркетинга.

Круг задач, решаемых геомаркетингом на уровне технологий гораздо шире того круга задач, которые решаются методами обычного маркетинга. Можно утверждать, что геомаркетинг решает практически все задачи, решаемые методами обычного маркетинга.

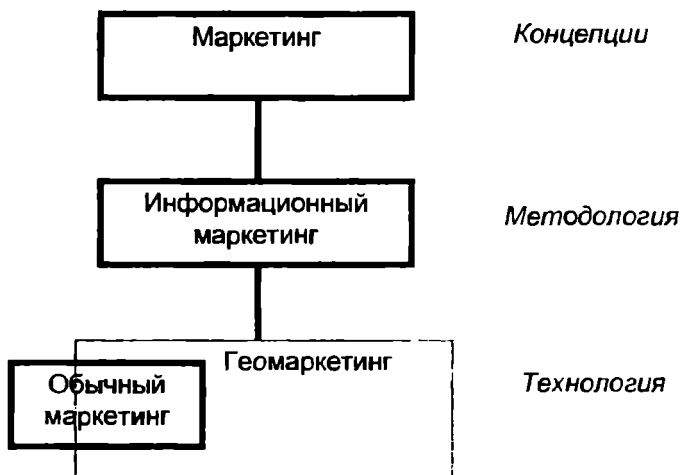


Рис.19.1. Соотношение обычного маркетинга и геомаркетинга

Однако использование геомаркетинга ограничено методами экономической целесообразности. Следует отметить тот факт, что как мощное интегрированное средство геомаркетинг не эффективен при решении частных небольших задач. В этих случаях целесообразней использовать методы обычного маркетинга.

Современные технологии обработки и анализа данных реализуются с использованием автоматизированных информационных систем. В маркетинге применяют маркетинговые информационные системы, в геомаркетинге - *геомаркетинговые информационные системы*.

Геомаркетинговые информационные системы возникли на основе интеграции геоинформационных систем и маркетинговых информационных систем. На рисунке 19.2 показано отношение между маркетинговой информационной системой, геоинформационной системой и геомаркетинговой информационной системой.

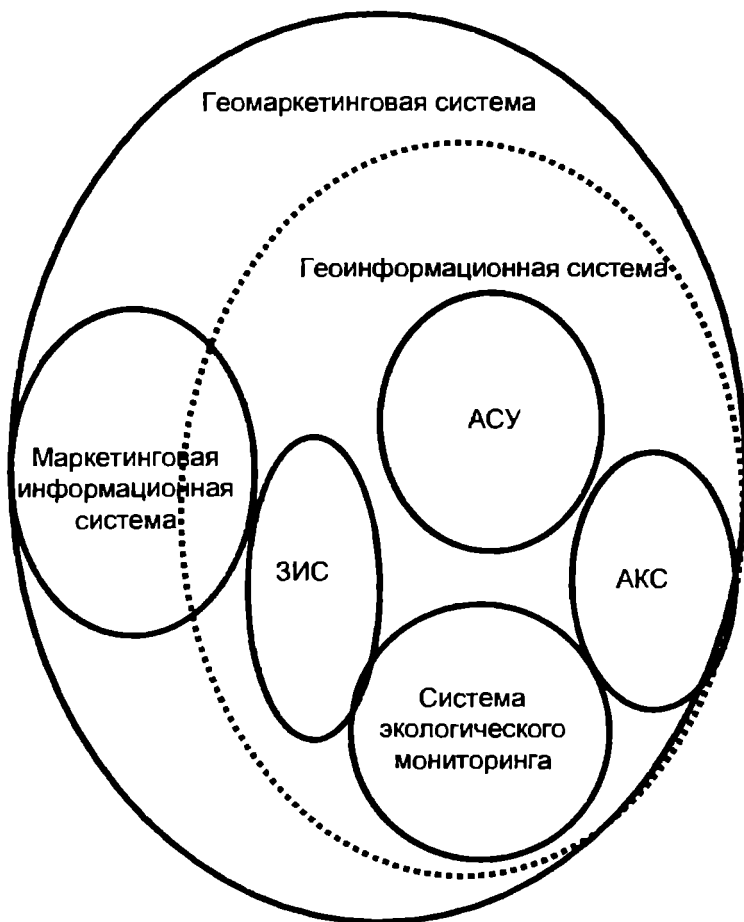


Рис. 19.2 Соотношение маркетинговой и геомаркетинговой информационных систем

Геомаркетинговая система позволяет решать комплексно ряд задач и включать в решение маркетинговых задач данные, которые ранее обрабатывались вне технологий и маркетинговых информационных систем. Такая интеграция данных и комплексная обработка относится к технологиям “Бизнеса в высоком темпе” CALS.

Применение геомаркетинговых систем и технологий целесообразно там, где возникает потребность работы с пространственно-локализованными данными или необходимо применять для поддержки принятия решений тематические карты с деловой графикой. Пространственная локализация может быть грубой или точной.

Точная пространственная локализация данных должна осуществляться при работе с картографической информацией или на участках, где необходим учет кривизны земной поверхности (20 x 20 км).

Следует отметить, что маркетинговые информационные системы могут работать с пространственно локализованными данными, но на основе классификации и специальных классификаторов (подобно тому как работают статистические информационные системы). Работа с классифицированными данными - это грубая привязка к точкам земной поверхности.

Геомаркетинговые информационные системы работают с пространственно-локализованными данными, включающими классификацию и позиционирование.

Классификация позволяет проводить качественный оперативный анализ. Позиционирование дает возможность проводить точный анализ. Оно является точной локализацией объектов.

Точная привязка данных дает три дополнительные возможности обработки данных в информационных технологиях.

1. Глобальная интеграция данных, позволяющая в совокупности комплексно изучать объекты и явления.

2. Возможность применения пространственного анализа объектов для выявления их свойств и отношений, не видимых при обычном анализе табличных или статистических данных

3. Применение визуальных методов представления и обработки статистической информации для поддержки принятия решений. Сюда следует отнести возможность обобщения (генерализации) однородных качественных явлений и характеристик.

Визуальный анализ данных на 2-3 порядка быстрее анализа табличных данных, особенно при контроле критических или аномальных ситуаций.



Перечисленные дополнительные возможности обработки определяют эффективность применения геомаркетинга как рыночной информационной технологии особенно при анализе пространственно-распределенных характеристик рынка: политическая ситуация, демографическая ситуация, экономическая ситуация, транспортные сети, туристические маршруты и т.д.

Анализ проектирования объектов большой протяженности (20 км и более) требует обязательного применения геоинформационных технологий, т.е. геомаркетинга.

## **19.1. Виды геомаркетинга**

### **19.1.1. Геомаркетинг мест**

Геомаркетинг как интегрированная технология характеризуется дифференциацией, т.е. выделением некоторых специальных технологий для решения специальных задач. К числу таких специальных технологий относится комплекс технологий, называемый геомаркетингом мест. В нем объектом маркетинга является некая пространственная единица.

**Геомаркетинг мест** - группа геомаркетинговых технологий, предпринимаемая с целью создания, поддержания или изменения отношений и/или поведения, касающихся конкретных мест. Выделяют следующие основные типы геомаркетинга мест: геомаркетинг жилья, геомаркетинг зон хозяйственной застройки, геомаркетинг инвестиций в земельную собственность, геомаркетинг мест отдыха, природоресурсный геомаркетинг.

**Геомаркетинг жилья** включает в себя застройку и/или активное предложение на продажу или внаем жилищ на одну семью, квартир и прочих жилых единиц.

**Геомаркетинг зон хозяйственной застройки** включает в себя хозяйственное освоение, продажу или сдачу в аренду участков или имущества типа заводов, магазинов, контор, складов и т. п. Крупные застройщики исследуют потребности фирм в земельных участках и предлагают комплексные решения проблем неподвижности типа создания промышленных зон, торговых центров, новых административных зданий. Целые города и страны заняты маркетингом своих территорий как благоприятных мест для капиталовложений в хозяйственную деятельность.

**Геомаркетинг инвестиций в земельную собственность** включает в себя обустройство и продажу земельных участков как объектов помещения капитала.

**Геомаркетинг мест отдыха и туризма** имеет целью привлече-

ние отдыхающих и туристов на курорты, в конкретные города, штаты, страны. Однако в ряде мест пытаются, наоборот, проводить демаркетинг.

Типичная схема применяемая в геомаркетинге этого типа. отдаленно напоминает карту, содержащую минимально необходимую информацию. Главная цель такой схемы изобразительная, а не метрическая.

### 19.1.2. Природоресурсный геомаркетинг

**Природоресурсный геомаркетинг** включает в себя хозяйственное освоение, продажу, привлечение инвестиций в природоресурсные региональные образования.

При организации и проведении этого маркетинга необходимо использовать понятия, определяющие региональное образование как некую совокупность, имеющую общие признаки и отличающуюся от других. Кроме того, необходимо определить природные ресурсы как особый объект геомаркетинга.

**Природные ресурсы** включают элементы природы и виды энергии, которые непосредственно участвуют в материальном производстве и составляют основу производимой продукции при данном уровне развития производительных сил.

Природные ресурсы находятся в тесном взаимодействии друг с другом, причем это взаимодействие усиливается по мере вовлечения их в производство. Природные ресурсы локализованы в регионах, имеют четкую привязку к месту ("место"), имеют характеристики объема, доступности ("тема"), требуют периодической актуализации ("время") и, таким образом, четко вписываются в геомаркетинговые данные и становятся объектами геинформационных технологий и геомаркетинга.

Под **природными ресурсами** понимают значительное число веществ, образовавшихся в поверхностной зоне земной коры - литосфере - и окружающей ее биосфере. Первые из них представлены минеральными образованиями, вторые - главным образом органическими.

Многообразие областей применения и аспектов рассмотрения природных ресурсов делает многоаспектной задачу их экономической классификации.

Наиболее простая классификация по области применения и характеру использования основана на делении ресурсов на энергетические и неэнергетические. **Энергетические ресурсы** - нефть, газ, уголь, торф, водные ресурсы, солнечная энергия, энергия приливов, отливов, ветра; **неэнергетические** - все остальные.

Другая классификация ресурсов подразделяет их на: топливно-энергетические, рудные, нерудные, водные.

Аспект истощаемости классифицирует природные ресурсы на две группы: **неисчерпаемые** (солнечная энергия, энергия приливов и отливов и т. д.); **исчерпаемые**, которые, в свою очередь, могут быть воспроизводимыми и невозпроизводимыми.

Каждая из таких классификаций по-своему позволяет определить место и значение различных природных ресурсов в производстве, а все в совокупности дают более полную характеристику природным ресурсам.

Ресурсы, используемые в сельском хозяйстве, делятся на ресурсы для выращивания полезных культурных растений и для непосредственного потребления.

Целью природоресурсного геомаркетинга является гармоничное развитие региона с сохранением его экологического баланса. При проведении природоресурсного геомаркетинга в зависимости от конкретной ситуации, складывающейся на рынке с точки зрения состояния спроса возможны следующие типы геомаркетинга: конверсионный, стимулирующий, развивающийся, ремаркетинг, демаркетинг

**Конверсионный геомаркетинг** - включает деятельность, связанную с преодолением негативных представлений о вторичных материальных ресурсах и разработкой такого плана, который способствовал бы зарождению спроса на соответствующие товары и услуги.

Как правило, состояние рынка на товары из отходов производства характеризуется отрицательным спросом. Это соответствует ситуации, когда часть потенциальных покупателей отвергает данный товар. Цель конверсионного геомаркетинга изменить ситуацию. Для этого используют такую его разновидность как стимулирующий геомаркетинг.

**Стимулирующий геомаркетинг** - это совокупность мер, направленная на преодоление негативного отношения (или безразличия) на товары или услуги. Он проводится при состоянии рынка с отсутствием спроса на товары или услуги. План геомаркетинга в этом случае должен изучить причины такого безразличия и определить мероприятия по его преодолению.

**Развивающийся геомаркетинг** - деятельность, связанная с формированием спроса на новые разработки технологических производств и новые виды товаров. Его используют тогда, когда состояние рынка, характеризуется потенциальным спросом под воз-

действием правовой ответственности. Например, требования экологической безопасности приводят к спросу на соответствующие средства ее обеспечения. Данный вид маркетинга направлен на снижение или исключение штрафных платежей за нарушение экологической безопасности, охрану окружающей среды и на уменьшение затрат на преодоление экологических катастроф.

**Ремаркетинг** - деятельность, направленная на оживление спроса на товары или услуги. Она обусловлена естественным снижением спроса на все виды товаров и услуг в период экономического спада их жизненного цикла. Цель ремаркетинга - оживление экономики и спроса при помощи новых возможностей маркетинга.

**Демаркетинг** - деятельность, направленная на снижение спроса на товары или услуги. Данный вид маркетинга применяется, когда спрос на некоторые товары значительно превышает предложения. Ситуацией применения демаркетинга может служить избыток туристов в зонах отдыха или угроза ажиатажного спроса.

### **19.1.3. Дополнительные виды геомаркетинга**

К специальным технологиям геомаркетинга относят геомаркетинг лиц, геомаркетинг организаций, общественный геомаркетинг, политический геомаркетинг.

**Геомаркетинг лиц** - деятельность, предпринимаемая для создания, поддержания или изменения позиций и/или поведения по отношению к конкретным лицам, проводимая в рамках определенных территориальных образований. Две наиболее распространенные формы этой деятельности - маркетинг знаменитостей и маркетинг политических кандидатов.

**Маркетинг организаций** - деятельность с целью создания, поддержания или изменения позиций и/или поведения целевых аудиторий по отношению к конкретным организациям. Маркетингом организаций традиционно занимаются отделы по организации общественного мнения. Это вытекает, в частности, из следующего определения технологии по организации общественного мнения:

**Технология организации общественного мнения** - технология, включающая следующие составляющие:

оценку общественного мнения к организации или лицу;

оценку деятельности лица или организации в соответствии с общественными интересами,

планирование и реализацию программы действий для завоевания понимания и восприятия его со стороны общественности.

Кроме того, существует **общественный маркетинг** направленный на претворение в жизнь программ, имеющих целью добить-

ся восприятия целевой группой (или целевыми группами) общественной идеи, движения или практики.

Сегментирование для определения регионально распределенных целевых групп - необходимое условие общественного геомаркетинга. Для достижения максимальной ответной реакции целевой группы изучают состав и особенности групп. На основе проведенных исследований разрабатывают замысел, проектируют коммуникации, приемы упрощающие понимание идеи и т.д. Более подробно остановимся на политическом геомаркетинге.

#### **19.1.4. Политический геомаркетинг**

**Политический геомаркетинг** – это деятельность, предпринимаемая для создания, поддержания или изменения позиций и/или поведения по отношению к политическим деятелям, партиям или организациям лицам. Данный вид геомаркетинга проводится в рамках определенных территориальных образований.

Как правило, объектами политического маркетинга являются не только отдельные лица или партии, но и политические идеи.

Политическим геомаркетингом занимаются специальные организации. Особенностью политического геомаркетинга является то, что в отличие от других видов маркетинга направленных на удовлетворение спроса в реальной продукции или услугах, политический геомаркетинг направлен на удовлетворение желаний, которые в дальнейшем в большинстве случаев не реализуются.

Политический геомаркетинг начинается с анализа электората и выделения целевых групп. Основой для выделения таких групп является геомаркетинговое исследование электората, первичный анализ информации, кластерный анализ и т.п.

Исследование электората и выделение групп осуществляется, как правило, многоаспектно (например, по социальным, демографическим, национальным, территориальным, религиозным – признакам). Для выявленных групп определяют идеи, вызывающие максимальную ответную реакцию группы.

По отношению к выдвигаемому политическому лидеру, партии и его политической программе осуществляют сегментацию электората.

Для успешного проведения политического геомаркетинга необходим учет конкуренции и обязательное противодействие конкурентам. По этой причине осуществляют сегментацию по отношению к продвигаемым политическим идеям или лицам и сегментацию по отношению к конкурентам. Результаты сегментирования могут дать различные оценки политической ситуации рынка.

Выявляют “окна” и “ниши” в различных типах сегментации по отношению к различным идеям.

**Окном** называют сегмент электората (политического рынка), которым пренебрегают наиболее близкие конкуренты в силу негативной реакции данной целевой группы на выдвигаемые идеи.

**Ниша** – это сегмент электората (свой избиратель), общественные интересы которого в максимальной степени соответствуют пропагандируемым идеям и лозунгам.

Анализ политической ситуации, как правило, осуществляют по территориальным образованиям: выборным округам, районам, субъектам федерации, экономическим регионам. Эта геомаркетинговая задача решается корректно только с применением геоинформационных технологий. Она не может быть решена с использованием технологий обычного маркетинга.

После геомаркетинговых исследований ситуации применяют так называемые выборные технологии. Эти технологии включают демаркетинг основных конкурентов и ремаркетинг собственных идей. На основе проведенных исследований разрабатывают, приемы упрощающие усвоение идеи, дополняют основные идеи другими для привлечения электората из других целевых групп и т.п.

При невозможности привлечения электорат из какой либо многочисленной целевой группы, работающей на конкурента, создают ему “фиктивного” противника. Этот противник ориентирован на ту же группу, что и основной конкурент и выступает примерно под теми же лозунгами. Но основной его задачей является отнятие голосов.

Включают технологии организации общественного мнения для искусственного завышения рейтинга нужного кандидата и понижения рейтинга его оппонента. Основным инструментом этих технологий являются средства массовой информации которые направлены на организацию:

- побуждения к единовременному действию,
- стремления изменить поведенческие привычки;
- изменение основополагающих представлений, если в этом есть необходимость.

Результаты геомаркетинговых исследований отображают в виде электронных или бумажных карт [28, 44, 48, 86, 106].

## Специальная геоинформатика

### 20. Выбор аппаратно-программной платформы

Выбор аппаратной платформы и конфигурации информационной системы представляет сложную задачу. Это связано определением рабочей нагрузки вычислительного комплекса в целом. Как правило, трудно с достаточной точностью предсказать саму нагрузку, особенно в случае, если система должна обслуживать несколько групп разнородных по своим потребностям пользователей. Обычно рабочая нагрузка определяется функциональным назначением системы [23, 24, 29, 37, 42, 57, 66, 93].

Решение задачи выбора конфигурации системы начинается с определения вида сервиса, который должен обеспечиваться системой, и определения того уровня сервиса, который может обеспечить выбираемая конфигурация.

Конфигурация аппаратных и программных средств определяется множеством разнородных, соединенных друг с другом компонентов системы.

Например, сети состоят из клиентов, серверов и сетевой инфраструктуры. Сетевая инфраструктура включает среду (часто нескольких типов) вместе с мостами, маршрутизаторами и системой сетевого управления, поддерживающей ее работу. В состав клиентских систем и серверов входят центральные процессоры, иерархия памяти, шин, периферийных устройств и программного обеспечения [25, 27, 66, 91].

Легко предсказать ограничения производительности системы по ее конфигурации, исходя из анализа наиболее слабых компонентов.

Поскольку современные комплексы почти всегда включают несколько работающих совместно систем, то точная оценка полной конфигурации требует ее рассмотрения как на макроскопическом уровне (уровне сети), так и на микроскопическом уровне (уровне компонент или подсистем).

Выбор аппаратной платформы и конфигурации определяется и рядом общих требований, предъявляемых к характеристикам современных вычислительных систем. К ним относятся:

- отношение стоимость/производительность;
- надежность и отказоустойчивость;
- масштабируемость;

- совместимость и мобильность программного обеспечения.

**Отношение стоимость/производительность.** Производительные средства стоят дорого, маломощные дешево. Приходится выбирать между этими двумя крайними направлениями, исходя из реальных финансовых ресурсов пользователя

Для сравнения различных компьютеров обычно применяются стандартные методики измерения производительности. Эти методики позволяют разработчикам и пользователям использовать полученные в результате испытаний количественные показатели для оценки тех или иных технических решений. Именно производительность и стоимость дают пользователю рациональную основу для решения вопроса, какой компьютер выбрать.

Перечислим и охарактеризуем параметры, необходимые для оценки качества компьютера.

**Надежность и отказоустойчивость.** Повышение надежности основано на принципе предотвращения неисправностей путем снижения интенсивности отказов и сбоев за счет применения электронных схем и компонентов с высокой и сверхвысокой степенью интеграции, снижения уровня помех, облегченных режимов работы схем, обеспечение тепловых режимов их работы, а также за счет совершенствования методов сборки аппаратуры.

**Отказоустойчивость** - это свойство вычислительной системы, которое обеспечивает ей возможность продолжения действий, заданных программой, после возникновения неисправностей. Введение отказоустойчивости требует избыточного аппаратного и программного обеспечения.

**Масштабируемость** вычислительной системы представляет собой возможность наращивания мощности процессоров, объемов оперативной и внешней памяти и других ресурсов вычислительной системы. Масштабируемость обеспечивается архитектурой и конструкцией компьютера, а также соответствующими средствами программного обеспечения.

**Совместимость и мобильность программного обеспечения.** Мобильность заключается в переносимости программного обеспечения. Концепция программной совместимости для архитектуры вычислительной системы состоит в создании такой архитектуры, которая была бы одинаковой (с точки зрения пользователя) для всех моделей системы независимо от цены и производительности каждой из них [24, 37, 39, 41, 42, 49].

В настоящее время одним из наиболее важных факторов, определяющих современные тенденции в развитии информационных



технологий, является ориентация компаний-поставщиков компьютерного оборудования на рынок прикладных программных средств.

Переход от гомогенных (однородных) сетей программно совместимых компьютеров к построению гетерогенных (неоднородных) сетей, включающих компьютеры разных фирм-производителей, превратил компьютерную сеть в средство интеграции информационных ресурсов.

Этот переход выдвинул ряд новых требований. Во-первых, такая вычислительная среда должна позволять гибко менять количество и состав аппаратных средств и программного обеспечения в соответствии с меняющимися требованиями решаемых задач. Во-вторых, она должна обеспечивать возможность запуска одних и тех же программных систем на различных аппаратных платформах, т.е. обеспечивать мобильность программного обеспечения. В третьих, эта среда должна гарантировать возможность применения одних и тех же человеко-машинных интерфейсов на всех компьютерах, входящих в неоднородную сеть.

В условиях жесткой конкуренции производителей аппаратных платформ и программного обеспечения сформировалась концепция открытых систем, представляющая собой совокупность стандартов на различные компоненты вычислительной среды, предназначенных для обеспечения мобильности программных средств в рамках неоднородной, распределенной вычислительной системы.

Одним из вариантов моделей открытой среды является модель OSE (Open System Environment), предложенная комитетом IEEE POSIX.

## **21. Классификация компьютеров, применяемых в геоинформационных технологиях**

### **21.1. Персональные компьютеры и рабочие станции**

Персональные компьютеры (ПК) появились в результате эволюции миникомпьютеров при переходе элементной базы машин с малой и средней степенью интеграции на большие и сверхбольшие интегральные схемы. ПК, благодаря своей низкой стоимости, очень быстро завоевали хорошие позиции на компьютерном рынке и создали предпосылки для разработки новых программных средств, ориентированных на конечного пользователя.

Миникомпьютеры стали прародителями и другого направления развития современных систем - 32-разрядных машин. Создание

RISC-процессоров и микросхем памяти емкостью более 1 Мбит привело к окончательному оформлению *настольных систем высокой производительности*, которые сегодня известны как рабочие станции.

В отличие от ПК, которые в начале ориентировались на самого широкого потребителя непрофессионала, рабочие станции ориентированы на профессиональных пользователей. Такая концепция привела к тому, что рабочие современные станции - это сбалансированные системы, в которых высокое быстродействие сочетается с большим объемом оперативной и внешней памяти, высокопроизводительными внутренними магистралями, высококачественной и быстродействующей графической подсистемой и разнообразными устройствами ввода/вывода. Это свойство выгодно отличает рабочие станции среднего и высокого класса от ПК и сегодня.

В то же время, быстрый рост производительности ПК на базе новейших микропроцессоров Intel в сочетании с резким снижением цен на эти изделия и развитием технологии локальных шин (VESA и PCI), позволяющей устранить многие "узкие места" в архитектуре ПК, делают современные персональные компьютеры весьма привлекательной альтернативой рабочим станциям.

В свою очередь производители рабочих станций создали изделия так называемого "начального уровня", которые по стоимостным характеристикам близки к высокопроизводительным ПК, но все еще сохраняют лидерство по производительности и возможностям наращивания.

Персональные компьютеры в настоящее время имеют достаточную производительность, а рабочие станции имеют программное обеспечение, способное выполнять большинство функций, которые стали ассоциироваться с понятием "персональной рабочей станции". Оба этих направления могут рассматриваться в качестве сетевого ресурса для информационных систем корпоративного уровня.

Среди других факторов, способствующих этому процессу, выделяют следующие:

- применение ПК стало более разнообразным, так как помимо обычных для этого класса систем текстовых процессоров, даже средний пользователь ПК может теперь работать сразу с несколькими прикладными пакетами, включая электронные таблицы, базы данных и высококачественную графику;
- адаптация графических пользовательских интерфейсов существенно увеличила требования пользователей ПК к соотношению производительность/стоимость. Несмотря на то,

что оболочка MS Windows может работать на моделях ПК 386SX с 2 Мбайтами оперативной памяти, реальные пользователи хотели бы использовать все преимущества подобных систем, включая возможность комбинирования и эффективного использования различных пакетов.

- Широкое распространение систем мультимедиа прямо зависит от возможности использования высокопроизводительных ПК и рабочих станций с адекватными аудио- и графическими средствами, и объемами оперативной и внешней памяти.
- Слишком высокая стоимость мейнфреймов и даже систем среднего класса помогла сместить многие разработки в область распределенных систем и систем клиент-сервер, которые многим представляются вполне оправданной по экономическим соображениям альтернативой. Эти системы прямо базируются на высоконадежных и мощных рабочих станциях и серверах.

## 21.2. Серверы

Прикладные многопользовательские ГИС-технологии, сетевые приложения и системы обслуживания коммуникаций, разработку программного обеспечения и обработку изображений требуют перехода к модели вычислений "клиент-сервер" и распределенной обработке.

В распределенной модели "клиент-сервер" часть работы выполняет сервер, а часть пользовательский компьютер.

Существует несколько типов серверов, ориентированных на разные применения: файл-сервер, сервер базы данных, принт-сервер, вычислительный сервер, сервер приложений. Тип сервера определяется видом ресурса, которым он владеет: файловая система, база данных, принтеры, процессоры или прикладные пакеты программ.

Другая классификация определяется масштабом сети, в которой используется сервер.: сервер рабочей группы, сервер отдела или сервер масштаба предприятия (корпоративный сервер). Эта классификация весьма условна. Например, размер группы может меняться в диапазоне от нескольких человек до нескольких сотен человек, а сервер отдела обслуживать от 20 до 150 пользователей.

Файловые серверы небольших рабочих групп (не более 20-30 человек) проще всего реализуются на платформе персональных компьютеров и программном обеспечении Novell NetWare. Файл-сервер, в данном случае, выполняет роль центрального хранилища данных.

Серверы прикладных систем и высокопроизводительные машины для среды "клиент-сервер" значительно отличаются требованиями к аппаратным и программным средствам.

Скорость процессора для серверов с интенсивным вводом/выводом не критична. Они должны быть оснащены достаточно мощными блоками питания для возможности установки дополнительных плат расширения и дисковых накопителей. Желательно применение устройства бесперебойного питания. Оперативная память обычно имеет объем не менее 32 Мбайт, что позволит операционной системе (например, NetWare) использовать большие дисковые кэши и увеличить производительность сервера. Как правило, для работы с многозадачными операционными системами такие серверы оснащаются интерфейсом SCSI (или Fast SCSI). Распределение данных по нескольким жестким дискам может значительно повысить производительность.

При наличии одного сегмента сети и 10-20 рабочих станций пиковая пропускная способность сервера ограничивается максимальной пропускной способностью сети. В этом случае замена процессоров или дисковых подсистем более мощными не увеличивают производительность, так как узким местом является сама сеть. Поэтому важно использовать хорошую плату сетевого интерфейса.

На базе многопроцессорных серверов обычно строятся также серверы баз данных крупных информационных систем, так как на них ложится основная нагрузка по обработке информационных запросов. Подобного рода серверы получили название суперсерверов.

Современные суперсерверы характеризуются:

- наличием двух или более центральных процессоров;
- многоуровневой шинной архитектурой, в которой запатентованная высокоскоростная системная шина связывает между собой несколько процессоров и оперативную память, а также множество стандартных шин ввода/вывода, размещенных в том же корпусе;
- поддержкой технологии дисковых массивов;
- поддержкой режима симметричной многопроцессорной обработки, которая позволяет распределять задания по нескольким центральным процессорам или режима асимметричной многопроцессорной обработки, которая допускает выделение процессоров для выполнения конкретных задач.

Как правило, суперсерверы работают под управлением операционных систем UNIX и Windows NT, которые обеспечивают многопоточковую многопроцессорную и многозадачную обработку.

Суперсерверы должны иметь достаточные возможности наращивания дискового пространства и вычислительной мощности, средства обеспечения надежности хранения данных и защиты от несанкционированного доступа. Кроме того, в условиях быстро растущей организации, важным условием является возможность наращивания и расширения уже существующей системы.

### 21.3. Мейнфреймы

Термин «мейнфрейм» употребляют как синоним понятия "большая универсальная ЭВМ". Мейнфреймы являются наиболее мощными (не считая суперкомпьютеров) вычислительными системами общего назначения, обеспечивающими непрерывный круглосуточный режим эксплуатации. Они включают один или несколько процессоров, каждый из которых, в свою очередь, может оснащаться векторными сопроцессорами (ускорителями операций с суперкомпьютерной производительностью).

Прогресс в области элементно-конструкторской базы позволил существенно сократить габариты этих устройств. Наряду со сверхмощными мейнфреймами, требующими организации двухконтурной водяной системы охлаждения, имеются менее мощные модели для охлаждения которых достаточно принудительной воздушной вентиляции и модели, построенные по блочно-модульному принципу и не требующие специальных помещений и кондиционеров.

Основными поставщиками мейнфреймов являются известные компьютерные компании IBM, Amdahl, ICL, Siemens Nixdorf и некоторые другие, но ведущая роль принадлежит безусловно компании IBM. Именно архитектура системы IBM/360, выпущенной в 1964 году, и ее последующие поколения стали образцом для подражания.

В архитектурном плане мейнфреймы представляют собой многопроцессорные системы, содержащие один или несколько центральных и периферийных процессоров с общей памятью, связанных между собой высокоскоростными магистралями передачи данных. При этом основная вычислительная нагрузка ложится на центральные процессоры, а периферийные процессоры (в терминологии IBM - селекторные, блок-мультиплексные, мультиплексные каналы и процессоры телеобработки) обеспечивают работу с широкой номенклатурой периферийных устройств.

Первоначально мейнфреймы ориентировались на централизованную модель вычислений, работали под управлением патентованных операционных систем и имели ограниченные возможности для объединения в единую систему оборудования различных фирм-поставщиков.

Повышенный интерес потребителей к открытым системам заставил поставщиков мейнфреймов существенно расширить возможности своих операционных систем в направлении совместимости. В настоящее время они демонстрирует свою "открытость", обеспечивая соответствие со спецификациями POSIX 1003.3, возможность использования протоколов межсоединений OSI и TCP/IP или предоставляя возможность работы на своих компьютерах под управлением операционной системы UNIX собственной разработки.

Стремительный рост производительности персональных компьютеров, рабочих станций и серверов создал тенденцию перехода с мейнфреймов на компьютеры менее дорогих классов: миникомпьютеры и многопроцессорные серверы. Эта тенденция получила название "разукрупнение" (downsizing).

Однако этот процесс в самое последнее время несколько замедлился. Основной причиной возрождения интереса к мейнфреймам эксперты считают сложность перехода к распределенной архитектуре клиент-сервер, которая оказалась выше, чем предполагалось. Кроме того, многие пользователи считают, что распределенная среда не обладает достаточной надежностью для наиболее ответственных приложений, которой обладают мейнфреймы.

Очевидно выбор центральной машины (сервера) для построения информационной системы предприятия возможен только после глубокого анализа проблем, условий и требований конкретного заказчика и долгосрочного прогнозирования развития этой системы.

Главным недостатком мейнфреймов в настоящее время остается относительно низкое соотношение производительность/стоимость.

Следует также помнить, что в мире существует огромная инсталлированная база мейнфреймов, на которой работают десятки тысяч прикладных программных систем. Отказаться от годами наработанного программного обеспечения просто не разумно. Эти системы, с одной стороны, позволят модернизировать существующие системы, обеспечив сокращение эксплуатационных расходов, с другой стороны, создадут новую базу для наиболее ответственных приложений.

#### **21.4. Кластерные архитектуры**

Двумя основными проблемами построения вычислительных систем для приложений, связанных с обработкой транзакций, управлением базами данных и обслуживанием телекоммуникаций, являются обеспечение высокой производительности и продолжительного функционирования систем. Наиболее эффективный способ достиже-

ния заданного уровня производительности - применение параллельных масштабируемых архитектур.

Задача обеспечения продолжительного функционирования системы имеет три составляющих: надежность, готовность и удобство обслуживания. Все эти три составляющих предполагают, в первую очередь, борьбу с неисправностями системы, порождаемыми отказами и сбоями в ее работе. Эта борьба ведется по всем трем направлениям, которые взаимосвязаны и применяются совместно.

Повышение надежности основано на принципе предотвращения неисправностей путем снижения интенсивности отказов и сбоев за счет применения электронных схем и компонентов с высокой и сверхвысокой степенью интеграции, снижения уровня помех, облегченных режимов работы схем, обеспечение тепловых режимов их работы, а также за счет совершенствования методов сборки аппаратуры.

Повышение уровня готовности предполагает подавление в определенных пределах влияния отказов и сбоев на работу системы с помощью средств контроля и коррекции ошибок, а также средств автоматического восстановления вычислительного процесса после проявления неисправности, включая аппаратную и программную избыточность, на основе которой реализуются различные варианты отказоустойчивых архитектур. Повышение готовности есть способ борьбы за снижение времени простоя системы. Основные эксплуатационные характеристики системы существенно зависят от удобства ее обслуживания, в частности от ремонтпригодности, контролепригодности и т.д.

В последние годы в литературе по вычислительной технике все чаще употребляется термин "системы высокой готовности" (High Availability Systems). Все типы систем высокой готовности имеют общую цель - минимизацию времени простоя.

Имеется два типа времени простоя компьютера: плановое и неплановое. Минимизация каждого из них требует различной стратегии и технологии. Плановое время простоя обычно включает время, принятое руководством, для проведения работ по модернизации системы и для ее обслуживания. Неплановое время простоя является результатом отказа системы или компонента. Хотя системы высокой готовности возможно больше ассоциируются с минимизацией неплановых простоев, они оказываются также полезными для уменьшения планового времени простоя.

Существует несколько типов систем высокой готовности, отличающиеся своими функциональными возможностями и стоимостью.

Стоимость систем высокой готовности превышает стоимость обычных систем в несколько раз. Вероятно поэтому наибольшее распространение в мире получили кластерные системы. Они обеспечивают достаточно высокий уровень готовности систем при относительно низких затратах.

Термин "кластеризация" в компьютерной технологии употребляется в различных значениях. В настоящей работе используется следующее определение.

**Кластер** - объединение компьютеров, представляющееся единым целым для операционной системы, системного программного обеспечения, прикладных программ и пользователей. Машины, кластеризованные таким способом могут при отказе одного процессора очень быстро перераспределить работу на другие процессоры внутри кластера. Это, возможно, наиболее важная задача многих поставщиков систем высокой готовности.

Первой концепцию кластерной системы анонсировала компания DEC, определив ее как группу объединенных между собой вычислительных машин, представляющих собой единый узел обработки информации. По существу VAX-кластер представляет собой слабовязанную многомашинную систему с общей внешней памятью, обеспечивающую единый механизм управления и администрирования. В настоящее время на смену VAX-кластерам приходят UNIX-кластеры. При этом VAX-кластеры предлагают проверенный набор решений, который устанавливает критерии для оценки подобных систем. VAX-кластер обладает рядом свойств:

**Разделение ресурсов.** Компьютеры VAX в кластере могут разделять доступ к общим ленточным и дисковым накопителям. Все компьютеры VAX в кластере могут обращаться к отдельным файлам лентных как к локальным.

**Высокая готовность.** Если происходит отказ одного из VAX-компьютеров, то задания его пользователей автоматически могут быть перенесены на другой компьютер кластера. Обычно при наличии в системе нескольких контролеров внешних накопителей при отказе одного из них другие контроллеры автоматически подхватывают его работу.

**Высокая пропускная способность.** Ряд прикладных систем могут пользоваться возможностью параллельного выполнения заданий на нескольких компьютерах кластера.

**Удобство обслуживания системы.** Общие базы данных обслуживаются с единственного места, а прикладные программы могут



инсталлироваться только однажды на общих дисках кластера и разделяться между всеми компьютерами кластера.

**Расширяемость.** Увеличение вычислительной мощности кластера достигается подключением к нему дополнительных VAX-компьютеров. Дополнительные накопители на магнитных дисках и магнитных лентах становятся доступными для всех компьютеров, входящих в кластер.

Работа любой кластерной системы определяется двумя главными компонентами: высокоскоростным механизмом связи процессоров между собой и системным программным обеспечением, которое дает клиентам прозрачный доступ к системному сервису.

В настоящее время широкое распространение получила также технология параллельных баз данных. Эта технология позволяет множеству процессоров разделять доступ к единственной базе данных. Распределение заданий по множеству процессорных ресурсов и параллельное их выполнение позволяет достичь более высокого уровня пропускной способности транзакций, поддерживать большее число одновременно работающих пользователей и ускорить выполнение сложных запросов. Существуют три различных типа архитектуры, которые поддерживают параллельные базы данных. Кратко их охарактеризуем.

1. Симметричная многопроцессорная архитектура с общей памятью (Shared Memory SMP Architecture). Эта архитектура поддерживает единую базу данных, работающую на многопроцессорном сервере под управлением одной операционной системы. Увеличение производительности таких систем обеспечивается наращиванием числа процессоров, устройств оперативной и внешней памяти.

2. Архитектура с общими (разделяемыми) дисками (Shared Disk Architecture). Это типичный случай построения кластерной системы. Эта архитектура поддерживает единую базу данных при работе с несколькими компьютерами, объединенными в кластер (обычно такие компьютеры называются узлами кластера), каждый из которых работает под управлением своей копии операционной системы. В таких системах все узлы разделяют доступ к общим дискам, на которых собственно и располагается единая база данных. Производительность таких систем может увеличиваться как путем наращивания числа процессоров и объемов оперативной памяти в каждом узле кластера, так и посредством увеличения количества самих узлов.

3. Архитектура без разделения ресурсов (Shared Nothing Architecture). Как и в архитектуре с общими дисками, в этой архитектуре поддерживается единый образ базы данных при работе с

несколькими компьютерами, работающими под управлением своих копий операционной системы. Однако в этой архитектуре каждый узел системы имеет собственную оперативную память и собственные диски, которые не разделяются между отдельными узлами системы. Практически в таких системах разделяется только общий коммуникационный канал между узлами системы. Производительность таких систем может увеличиваться путем добавления процессоров, объемов оперативной и внешней (дисковой) памяти в каждом узле, а также путем наращивания количества таких узлов.

Таким образом, среда для работы параллельной базы данных обладает двумя важными свойствами: высокой готовностью и высокой производительностью. В случае кластерной организации несколько компьютеров или узлов кластера работают с единой базой данных. В случае отказа одного из таких узлов, оставшиеся узлы могут взять на себя задания, выполнявшиеся на отказавшем узле, не останавливая общий процесс работы с базой данных. Поскольку логически в каждом узле системы имеется образ базы данных, доступ к базе данных будет обеспечиваться до тех пор, пока в системе имеется по крайней мере один исправный узел. Производительность системы легко масштабируется, т.е. добавление дополнительных процессоров, объемов оперативной и дисковой памяти и новых узлов в систему может выполняться в любое время, когда это действительно требуется.

Параллельные базы данных находят широкое применение в системах обработки транзакций в режиме on-line, системах поддержки принятия решений и часто используются при работе с критически важными для работы предприятий и организаций приложениями, которые эксплуатируются по 24 часа в сутки.

## **22. Кодирование и элементы теории информации.**

### **22.1. Постановка вопроса.**

При автоматизированном сборе данных об объектах и явлениях возникает необходимость их кодирования. Это обусловлено тем, что информация для целей ее машинного хранения и обработки всегда должна быть представлена в строго определенной форме [24, 29, 39, 49, 57, 66].

На низшем уровне информационного развития наиболее распространенной формой представления информации являются тек-

сты. С точки зрения информатики текст является конечной последовательностью знаков.

Существует много различных систем представления информации, базирующихся на последовательностях знаков. Различные системы представления по-разному удобны для машинной обработки.

Исследование простых и экономичных представлений информации с помощью последовательностей знаков приводят к необходимости изучения вопросов кодирования информации и кодам.

**Кодировка (или код)** представляет собой систему, позволяющую осуществлять переход от одной знаковой системы представления информации к другому представлению той же информации также в виде знаков и их последовательностей. На этом этапе следует подчеркнуть отличие кодировки от шифрования, применяемого при защите информации. При шифровании последовательность знаков сознательно меняется и допускается неоднозначность представления исходного текста в зашифрованном.

При выборе способа кодировки в первую очередь учитываются два фактора: экономичность представления и обработки, а также мера надежности от ошибок.

Из соображений эффективности используют по возможности короткие кодовые слова, с тем чтобы представление информации в виде кодов было по возможности компактным, наглядным и более дешевым. Кроме того, обработка информации в выбранной системе представления должна быть простой и экономичной.

Длина слова связана с размером алфавита. Так в системе компьютерного кодирования двоичные слова наиболее длинны, а шестнадцатеричные являются самыми короткими.

С другой стороны, если в процессе передачи или обработки информации в кодовых словах появляются случайные незначительные ошибки, то желательно такие «испорченные» кодовые слова выявлять и декодировать.

При установке средней частоты (вероятности) появления кодируемой информации равной средней частоте появления отдельных знаков естественно выбрать кодировку так, чтобы средняя длина кода была по возможности меньше. Отсюда следует правило: при известной вероятности помех необходимо выбрать кодировку такой, чтобы вероятность не выявления помех была достаточно малой.

## 22.2. Коды и кодирование

Назовем конечное множество  $A$  знаков *запасом знаков*. Если это множество линейно упорядочено, то его будем называть также *алфавитом*. Особенным элементарным запасом знаков является

множество  $B$ :

$$B = \{L, 0\}.$$

Это множество содержит всего два знака  $L$  или  $0$ . Элемент множества  $B$  называют *двоичным знаком* или *битом* (Bit – от Binary Digit).

Множество конечных последовательностей знаков над запасом знаков  $A^*$  называется также *множеством слов над  $A$* .

При наличии множества слов  $B^* = \{L, 0\}^*$  говорят о *двоичных словах*. Элементы этого множества называют также  *$n$ -битовыми словами* или *двоичными словами длины  $n$* . Множество  $n$ -битовых слов часто само снова используется в качестве запаса знаков.

Кодами или кодировкой называют отображения между запасами знаков  $A$  и  $B$ . Код " $c$ " есть отображение

$$c: A \rightarrow B$$

или, в случае кодировки целых слов, отображение

$$c: A^* \rightarrow B^*$$

Специальные случаи кодировки предусматривают представление не для всех слов из  $A^*$ . Такие кодировки соответствуют частичным отображениям или отображениям вида

$$c: A' \rightarrow B' \quad (A' \in A^* \text{ и } B' \in B^*)$$

При частичном отображении может выбираться число

$$n \in N$$

для ограничения множества знаков

$$A' = A^n \text{ и } B' = B^n$$

Оно определяет отображение ограниченного количества слов или знаков множества  $A$  на множество  $B$ . Если запас знаков  $A$  состоит из отдельных литер, то говорится также о *шифровке* и множество образов называем *шифрами*.

В общем случае предполагается, что кодировка является инъективным отображением, что означает: различные знаки и слова отображаются на разные кодовые слова. Это называют однозначным соответствием. Тем самым каждая кодировка

$$c: A \rightarrow B$$

обратима на ее прообраз. Обратное отображение " $d$ "

$$d: \{c(a): a \in A\} \rightarrow A$$

для отображения кодировки "с" называем *декодировкой* или *дешифровкой*. Тогда справедливо следующее:

$$d \{c(a)\} = a$$

### 22.3. Коды постоянной длины

Коды постоянной длины это коды, которые каждый кодируемый знак отображают на (двоичное) слово одинаковой длины. Таким кодировкам соответствуют отображения вида

$$c:A \rightarrow B^n$$

Такая кодировка диктуется техническими характеристиками (например, фиксированная длина регистров вычислительной машины) вычислительных систем. Данный подход использует для представления кодируемого слова постоянное количество бит. В настоящее время в информатике и геоинформатике наиболее применяются коды постоянной длины. Рассмотрим ряд простых кодировок постоянной долины и обсудим их свойства.

Пусть  $a, b$  - двоичные слова одинаковой длины; число позиций, в которых различаются слова  $a$  и  $b$ , назовем *расстоянием Хэмминга*.

$$Ham\_dis: B^n \times A^n \rightarrow N$$

Таким образом, расстоянию Хэмминга математически соответствует отображение

$$c:A \rightarrow B^n$$

Расстояние отображения кодировки

$$Ham\_dis(c) = \min\{Ham\_dis(c(a), c(b)) : a, b \in A^n, a \neq b\}$$

для длины кодов  $n$  определяется через минимальное расстояние Хэмминга между двумя различными кодами.

Большее расстояние кода допускает распознавание ошибок, а часто даже их исправление, однако при этом необходимы более длинные коды и тем самым дополнительные затраты на представление. Это ведет к понятию избыточности.

Для алфавита  $A$  (т. е. для набора знаков  $A$ , на котором задан линейный порядок) код постоянной длины называется кодом Грэя (или *одношаговым кодом*), если коды двух последовательных знаков из  $A$  различаются точно в одной позиции (в одном бите).

Пример (4-битовый код Грэя для десятичных цифр). Четырех-битовый код Грэя для десятичных цифр соответствует кодировке

$$c : \{0, \dots, 9\} \rightarrow B^4$$

Что в развернутом виде показано в таблице

Таблица кода Грэя

Знак	Кодировка	Знак	Кодировка
1	0001	6	0110
2	0010	7	0111
3	0011	8	1000
4	0100	9	1001
5	0101	10	1010

Если в кодах Грэя кодировки первого и последнего знаков отличаются тоже только в одной позиции (как в вышеприведенном примере), то говорят о *циклическом коде Грэя*.

Большие расстояния Хэмминга могут быть достигнуты, например, увеличением длины кодовых слов для дополнительных знаков.

#### 22.4. Коды переменной длины

С кодами переменной длины технически труднее оперировать, чем с кодами постоянной длины. Из-за того, что в современных компьютерах используются слова определенной длины, такие коды в практике машинной обработки информации используются редко.

Примеры (коды переменной длины).

Пример 1. Код Морзе в двоичном кодировании

Код Морзе строится из трех знаков, которые соответствуют «короткой», «длинной» передаче (точки и тире) и пробелу («пауза»). Двоичная кодировка

$c : \{., -, \text{«пауза»}\} \rightarrow \{0L, 0LLL, 000\}$  определяется следующим образом:

$c(.) = 0L$ ,  $c(-) = 0LLL$ ,  $c(\text{«пауза»}) = 000$ .

Пример 2. Код для телефонных систем

Телефонная система использует код для цифр диска набора номера. Такой код, приведенный в таблице кода телефонных систем, порождается при наборе соответствующей цифры и передается в промежуточный узел.

Таблица. Код телефонных систем

Цифра	Код
1.	LO
2.	LLO
3.	LLLO
4.	LLLLO
5.	LLLLLO
6.	LLLLLLO
7.	LLLLLLLO
8.	LLLLLLLLLO
9.	LLLLLLLLLLO
10.	LLLLLLLLLLLO

У кодов переменной длины есть одно важное преимущество. При наличии данных о средней частоте появления знаков из  $A$  для часто встречающихся знаков можно выбрать короткие кодовые слова, а для редко встречающихся знаков - длинные, и тем самым обеспечить небольшую среднюю длину кодовых слов.

Трудности использования кодов переменной длины обусловлены тем, что, кодированию подлежат не отдельные знаки, а последовательности знаков. Если кодировка последовательности знаков производится конкатенацией кодов отдельных знаков, то при кодах переменной длины труднее распознать стыки знаковых кодов.

## 22.5. Последовательное и параллельное кодирование

При необходимости кодирования отдельных знаков и слов из  $A$  и передачи закодированной информации, то - исходя из кодирования отдельных знаков из  $A$  - могут использоваться две принципиально различные возможности.

1. Последовательная кодировка слов. Для заданной кодировки

$$c : A \longrightarrow B^*$$

отдельных знаков рассмотрим ее расширение на слова над  $A$ . Это значит, что рассматривается кодировка  $c^*$  над  $A^*$ , индуцированная кодировкой  $c$ :

$$c^* : A^* \longrightarrow B^n$$

и определяемая следующим образом:

$$c^*(\varepsilon) = \varepsilon$$

$$c^*(\langle a_1 \dots a_n \rangle) = c(a_1) \oplus \dots \oplus c(a_n)$$

$\oplus$  - символ конкатенции

Таким образом, при последовательной кодировке слов кодируются слова путем конкатенации кодов отдельных знаков слов.

**2 Параллельная кодировка слов.** Для заданной кодировки с постоянной длиной кодов

$$c: A \rightarrow B^n$$

$$c'(\varepsilon) = \varepsilon$$

$$c'(\langle a_1 \dots a_n \rangle) = \langle c(a_1) \dots c(a_n) \rangle$$

рассматривается индуцированное отображение  $c'$ :

Итак, при последовательной кодировке слов коды отдельных знаков подлежащего кодированию слова конкатенируются, в то время как при параллельном кодировании слов сохраняется структура кодируемого слова.

## 22.6. Оптимальность кодов и разрешающая информация

Особая точка зрения при выборе кодовой функции появляется в случае, когда для заданного множества знаков  $A$  известна определенная, не зависящая от времени, вероятность (средняя частота) вхождения знаков из  $A$  в рассматриваемую последовательность из  $A^*$ . Источник последовательности знаков с этими свойствами (каждому моменту времени для новых подлежащих пересылке знаков соответствуют вероятности, точно равные заданной средней частоте) называется *стохастическим* или *шенноновским источником сообщений*.

**Пример** (стохастический источник сообщений). Стохастический источник сообщений для заданного алфавита  $A$  подается, например, следующим образом. Возьмем любое количество шаров и каждый шар пометим знаком из алфавита  $A$  (при этом один и тот же знак может встретиться на нескольких шарах). Положим помеченные шары в коробку и перемешаем их. Затем из коробки случайным образом возьмем шар, запишем содержащийся на нем знак, бросим шар обратно в коробку и перемешаем шары. После этого опять случайным образом возьмем шар из коробки и т. д.

Таким способом мы пополним стохастический источник сообщений. Средняя частота (вероятность) появления какого-либо знака в порождаемой последовательности соответствует здесь отношению



числа шаров, помеченных данным знаком, к общему числу шаров.

Обратим внимание на то, что последовательности знаков, встречающиеся в естественных языках, не являются стохастическими. Вероятность того, что определенная буква следует за определенной заданной буквой в немецком, например, языке не является независимой от заданной буквы. В таком случае мы говорим о *контактной частоте*.

В дальнейшем мы опишем способ кодирования последовательностей знаков стохастического источника сообщений для получения минимально короткой средней длины кода сообщения. С использованием этого способа с успехом могут кодироваться сообщения не стохастических источников (например, сообщения на естественном языке).

Прежде чем обратиться к вопросу кодирования, обсудим понятие информационного содержания знака, который встречается в последовательности знаков стохастического источника сообщений.

Можно *разрешающую информацию* или *информационное содержание* знака примерно отождествить с обратной величиной частоты появления знака. Это наглядно соответствует нашему представлению об информационном содержании сообщений:

- сообщение о фактическом наступлении почти всегда наступающего события имеет небольшое информационное содержание;
- сообщения о фактическом наступлении редкого события, напротив, имеют высокое информационное содержание (другими словами, чем реже поступает сообщение, тем выше его информационное содержание).

## 22.7. Надежность передачи сообщений

Концепция избыточности может использоваться для уменьшения ошибочных передач при общении между людьми и при техническом обмене сообщениями, чтобы предохранить закодированную информацию от помех и ошибок.

Рассмотрим типичную ситуацию при передаче сообщений. От источника передаются параллельно сообщения через канал связи к некоторому приемнику (адресату) (рис. 22.1.).



Рис. 22.1. Схема передачи информации

Из-за ошибок передачи (помех) передаваемые сообщения могут быть искажены. В простейшем случае предположим, что каждый бит с определенной вероятностью подвержен искажению, причем независимо от передачи предшествующего знака. Говорится здесь о дискретном канале без памяти, и в дальнейшем ограничимся рассмотрением этого простого случая.

При так называемом двоичном канале из-за ошибок передачи всегда возникают тоже двоичные знаки. Оба знака  $O$  и  $L$  имеют вероятности ошибок соответственно  $P_O$  и  $P_L$ .

Специальными являются случаи односторонних помех, при котором имеет место  $P_O = 0$  или  $P_L = 0$  и случай симметричных помех, когда  $P_O = P_L$ . В случае симметричных помех с вероятностью  $p = 1/2$  для каждого передаваемого знака вероятности ошибочной и корректной передач одинаковы. Принятый знак не оказывает никакого влияния на передаваемый знак.

При канале с потерей знаков из-за ошибок передачи появляются испорченные знаки которые распознаются как искаженные знаки и могут быть представлены знаком искажения  $I$ . Такого рода знак приходит, например, при передачах, которые нарушаются через шум. Таким образом, канал в качестве входа принимает слова над множеством знаков  $\{L, O\}$  и выдает слова над множеством знаков  $\{L, O, I\}$ .

## 22.8. Надежность кодов

Важным приложением информатики является *распознавание ошибок и коррекция ошибок* при передаче сообщений. В связи с этим говорится о *надежности кодов*. Для повышения надежности используют избыточность при кодировании. Дополнительная избыточность возникает в тех случаях, когда при кодах постоянной длины используются не все комбинации битов в процессе кодирования. Подмножество двоичных слов, используемое для кодирования, может быть выбрано так, чтобы:

- (1) вероятность полнить неверный знак была по возможности

мала;

(2) вероятность обнаружения ошибки была по возможности больше;

(3) вероятность того, что появившаяся ошибка исправима, была возможно больше.

При двоичном кодировании в случае односторонних помех при  $m$ -из- $n$  кодах каждая помеха ведет к слову, которое не является кодовым словом, и потому может быть обнаружена.

При симметричных помехах для обнаружения помех требуется увеличение расстояния Хэмминга. Мы получаем следующую, легко доказываемую лемму.

**Лемма (обнаружение ошибок).** Если код имеет расстояние Хэмминга  $h$ , то все ошибки, которые встречаются меньше чем в  $h$  битах, могут быть обнаружены.

При предположениях, сформулированных в этой лемме, справедливо, что все ошибки, появившиеся меньше чем в  $h/2$  битах, могут быть успешно устранены, если для коррекции обнаруженной ошибки применяется кодовое слово с наименьшим расстоянием Хэмминга  $Hd$ . Получаются следующие простые следствия: если  $Hd = 2$ , то ошибочный бит может быть обнаружен; если  $Hd = 3$ , то могут быть обнаружены два ошибочных бита, а один ошибочный бит может быть успешно исправлен; корректировка передачи с двумя ошибочными битами ведет к неопределенному, а следовательно, некорректное результату.

Чтобы достичь исчерпывающей сохранности кодов, заданные коды часто снабжаются дополнительными битами: «битами четности» или «контрольными битами».

**Пример (код флексорайта).** К каждому кодовому слову добавляется контрольный бит. Его значение берется таким, чтобы общее число  $L$  в коде было четным. Если в кодовом слове с контрольным битом испорчен только 1 бит, то контрольный бит для возникшего слова некорректен и тем самым ошибка обнаруживаема.

Подчеркнем, что адекватный выбор и подходящее число контрольных битов существенно определяются характером возможным помех.

## 22.9. Надежность передачи

Рассмотрим двоичный канал для передачи двоичных слов, о которых (и их избыточности) мы ничего не знаем. Пусть в канале имеют место симметричные помехи. Это означает, что имеется фиксированная вероятность ошибок, с которой выход в одном знаке отличается от входа. Допустим, далее, принимая во внимание скорость

передачи (мощность канала), что источник должен передать определенное число  $S_0$  знаков, а канал может передать  $S_1$  знаков в единицу времени. Тогда  $R = S_0/S_1$  назовем *темпом источника*. Если  $R < 1$ , то возможно - с помощью соответствующей кодировки входа для передачи - достичь еще дополнительной избыточности и тем самым устойчивости к ошибкам. Таким образом, передача сообщения следует схеме, приведенной на рис. 22.2.

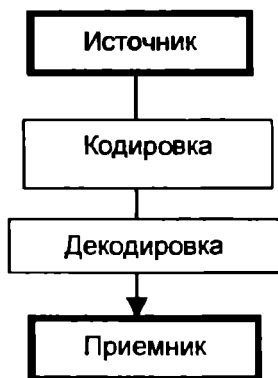


Рис.22.2. Схема передачи кодируемого сообщения

В случае темпа источника  $R < 1$  можно было бы использовать избыток мощности канала для того, чтобы с помощью неоднократной передачи информации раньше обнаруживать ошибки передачи и надежно их исправлять.

**Пример** (надежность через избыточность). Если, например, имеет место  $R = 1/3$ , то канал может передать втрое больше знаков, чем их вырабатывает источник. Каждый знак может быть троекратно закодирован перед передачей и в конце концов должен быть снова декодирован:

пересылаемое слово    **LOLOO,**  
 кодировка                **LLLOOOLLLOOOOOO.**

Если предположим, например, что вследствие шума биты с номерами ] 5, 6, 12 и 13 будут искажены (норма ошибок  $1/3$ ), то адресат получит последовательность знаков

**LOLOLLLLLOOLLOO.**

Лучшая стратегия декодирования с исправлением ошибок получается с помощью «вотума большинства» для тройки передавае-

мых битов. В рассматриваемом примере это приводит к получению адресатом следующей последовательности знаков:

LLLOO.

Некорректно будет устранена только ошибка в передаче второго бита эта ошибка и останется в декодированном сообщении. Норма ошибок тогда задается формулой

$$3p^2 - 2p^3$$

так что для малой вероятности ошибки  $p$  это значительно лучше, чем вероятность ошибки  $p$  при однократной передаче знака. Ошибка достигает адресата только в том случае, когда будут искажены 2 или 3 бита в кодировке одного знака. Если темп источника больше  $\frac{1}{2}$ , то стратегия неоднократной передачи непосредственно невозможна. Тогда можно работать с дополнительными битами, например битами четности, которые, однако, должны выбираться очень тщательно. Следующая стратегия битов четности называется *кодом Хэмминга*: если имеет место  $R = n/(n+1)$ , то возможен только один дополнительный бит для  $n$ -битового кода. Для передачи  $n$ -битового двоичного слова

$$\langle x_1 \dots x_n \rangle$$

выбирается бит  $x_{n+1}$  таким образом, что с  $W(L) = 1$  и  $W(0) = 0$  справедливо

$$(w(x_1) + \dots + w(x_n) + w(x_{n+1})) \bmod 2 = 0$$

Это значит, что каждое двоичное слово дополняется еще одним знаком так, чтобы число знаков  $L$  во всем кодовом слове было четным. Расстояние Хэмминга для  $(n+1, n)$ -кода Хэмминга, определяемого следующим образом:

$$c(\langle x_1 \dots x_n \rangle) = \langle x_1 \dots x_n x_{n+1} \rangle$$

равно по меньшей мере двум. Для такого кода единичная ошибка обнаруживается, но не устраняется.

Если, например,  $R = 4/7$  и пересылаются слова длины 4, то можно свободно выбирать 3 бита четности. Если должно передаваться слово

$$\langle x_0 \dots x_3 \rangle$$

по при  $(7, 4)$ -коде Хэмминга биты  $x_4, x_5, x_6$  могут быть выбраны в качестве битов четности.

Определим, например

$$w(x_4) = (w(x_1) + w(x_2) + w(x_3)) \bmod 2$$

$$w(x_5) = (w(x_0) + w(x_2) + w(x_3)) \bmod 2$$

$$w(x_6) = (w(x_0) + w(x_1) + w(x_2)) \bmod 2$$

Искажения битов  $x_4, x_5, x_6$  соответственно  $x_0, x_1, x_2, x_3$  при передаче не могут быть распознаны как ошибки. Расстояние Хэмминга для (7, 4)-кода Хэмминга равно трем. При (7, 4)-коде Хэмминга может быть исправлена единичная ошибка в кодовом слове.

При декодировании (в предположении, что имеется не более одной ошибки) можно поступать следующим образом. Адресат при пересылке двоичного слова  $x = \langle x_0 \dots x_6 \rangle$  получает, возможно, испорченное слово  $y$ . Если  $y$  не является кодовым словом, т. е.  $y$  не удовлетворяет приведенным выше равенствам, то мы считаем, что имеет место ошибка передачи и в качестве передаваемого слова выбираем то слово из кода Хэмминга, которое отличается от слова  $y$  только в одном бите. Если такого двоичного слова не существует, то исходное передаваемое слово однозначно не восстановимо.

Итак, наряду с вероятностью ошибки  $P$  для канала на передаваемый бит при темпе источника  $R < 1$  при использовании избыточности мы получаем меньшую вероятность ошибки  $P_e < p$  на подлежащий передаче знак. Если имеет место  $R > 1$ , то мы ожидаем  $P_e > p$

Если темп источника  $R$  больше единицы, то надо передать больше знаков, чем может справиться канал, и тогда не все биты источника могут быть переданы. Простейший метод декодирования при этом заключается в том, что адресат для непереданных битов бросает монету и таким способом угадывает значения непереданных битов. Вероятность ошибки  $P_e$  на бит в подлежащем передаче слове поддается тогда из следующей формулы:

$$P_e = p/R + (R-1)/(2R).$$

Лучшая мера корректно переданных битов поддается путем их упаковки в передаваемые блоки. Проиллюстрируем это на примере. Для  $R = 3$  источник делит сообщение на блоки трехразрядных двоичных слов и пересылает значение большинства. Адресат утраивает каждый переданный бит. Тогда мы получаем в совокупности вероятность ошибки

$$P_e = (2p+1)/4.$$

При вероятности ошибки  $P = 0$  примерно  $\frac{1}{4}$  знаков передаются корректно.

Описанная выше идея улучшения значения  $P_e$  может быть применена и в других случаях, при условии, что для темпа источника справедливо  $R > 0$ . Всегда можно пытаться многие биты передаваемого слова уплотнить в 1 бит (или в меньшее число бит), затем передать это и добавить целиком к подлежащему передаче сообщению. Это приводит к вопросу: насколько при заданных темпе источника  $R$  и вероятности  $P$  ошибки в бите можно улучшить с помощью кодирования вероятность ошибки  $P_e$  при передаче и подлежащий передаче бит? Принципиально можно сказать:

в случае  $R = 1$  определенно  $p_e = p$  является достижимым оптимумом, в случае  $R < 1$  определенно  $p_e < p$ , а в случае  $R > 1$  определенно  $p_e > p$ .

Точное высказывание относительно достижимой вероятности ошибки при заданных темпе источника и вероятности помехи может быть сделано следующим образом. Введем еще *двоичную функцию энтропии*  $H_2(p)$ , определяемую формулой

$$H_2(p) = p \cdot \lg(1/p) + (1 - p) \lg(1/(1-p))$$

В частности, справедливо  $H_2(0) = H_2(1) = 0$ .

Меру для достижимой надежности передачи дает мощность канала, которая указывает, какая вероятность ошибки при заданных темпе источника и вероятности ошибки на бит может быть достигнута в лучшем случае. Мощность  $C$  определяется

$$C(p) = 1 - H_2(p).$$

## 23. Представление пространственных объектов в виде множеств

Множество встречается во многих приложениях. На практике множества являются конечными, хотя и могут иметь очень большую мощность. Эффективность алгоритмов, работающих с множеством данных, зависит от того, могут ли быть быстро выполнены операции доступа к множествам.

Именно поэтому для представления больших множеств часто выбирают такие структуры данных, на которых операции доступа могут быть реализованы эффективно.

Если рассматриваются маленькие множества и необходимы операции пересечения и объединения, то можно порекомендовать представление множеств в виде битовых векторов. При этом под-

множества основного множества представляются векторами битов, длина которых соответствует мощности основного множества. В том случае, когда какой-либо элемент принадлежит множеству, то в соответствующий бит заносится значение `true`, в противном случае — значение `false`.

Такое представление плохо подходит для большого основного множества, поскольку теряется много памяти, особенно в том случае, когда должны представляться подмножества с относительно малой мощностью. В подобных случаях используется методика хэширования, которая будет рассмотрен в следующем разделе.

### 23.1. Вычислительная структура множеств с доступом по ключу

В базах данных часто возникает необходимость организации структуры данных, доступ к которым осуществляется с помощью ключей. Пусть дано множество ключей с помощью типа `sort key`.

```
fct key=(data) key.
```

Выше приведенная функция осуществляет доступ к порциям данных типа `data`, содержащих в себе компоненту типа `key`.

При этом предполагается, что каждая порция данных идентифицируется своим ключом. Тогда обращаться к нужной порции данных можно путем задания ключа. Необходимо найти тип `store`, который позволяет заносить порции данных и получить эффективный доступ к ним. При этом используются следующие функции:

```
fct emptystore=store,
fct get=(store, key) data,
fct insert=(store, data) store,
fct delete=(store, key) store.
```

Способ действия этих функций можно представить следующими равенствами:

```
k = key(d) ⇒ get(insert(s, d), k) = d
k ≠ key(d) ⇒ get(insert(s, d), k) = get(s, k),
delete(emptystore, k) = emptystore,
k = key(d) ⇒ delete(insert(s, d), k) = delete(s, k),
k ≠ key(d) ⇒ delete(insert(s, d), k) = insert(delete(s, k), d).
```

Ниже будут представлены структуры данных для представления больших множеств данных, которые позволяют эффективно реализовывать указанные выше операции.

### 23.2. Метод хэширования

Метод хэширования позволяет хранить множество элементов в



линейном массиве длиной  $z$ . Для этого нужна функция расстановки («рассыпания»):

$h: \text{key} \rightarrow \{0:z-1\}$ ,

которая каждый элемент типа `key` отображает на индекс в множестве  $\{0:z-1\}$ . Эта функция устанавливает, под каким индексом будет храниться данный элемент в массиве. Используем  $h(m)$  в качестве индекса (также называемого ключом) для запоминания элемента данных в массиве

`sort store = {0:z-1} array data a.`

Как правило, число элементов типа `key` значительно больше, чем  $z$ . Тогда функция  $h$  наверняка неинъективна. Возможно хранение элемента  $b$  с ключом  $m$  в массиве  $a$  под индексом  $h(m)$ . Получаются следующие реализации функций:

`fct emptystore = store: emptyarray,`

`fct get = (store a, key k) data: a[h(k)]`

`fct insert = (store a, data d) store: update(a, h(key(d)), d),`

`fct delete = (store a, key k) store: update(a, h(k), empty).`

Здесь предполагается, что `empty` обозначает элемент типа `data`, который играет роль держателя места. Функции работают корректно, до тех пор, пока для всех встречающихся ключей значения функции расстановки различны. Возникает проблема, когда нужно запоминать два различных элемента с ключами  $m_1$  и  $m_2$ , и при этом оказывается  $h(m_1)=h(m_2)$ . В этом случае говорят о *коллизии*.

Для использования метода хэширования нужно решить следующие проблемы:

- определения величины массива и тем самым числа  $z$  значений индексов,
- выбор функции расстановки  $h$ ,
- определение способа разрешения коллизий.

При этом имеются экспериментальные данные относительно того, сколь большим должен быть выбран массива, чтобы, с одной стороны, вероятность коллизий не была слишком велика и, с другой стороны, не пропадало слишком много памяти, если заняты не все позиции массива.

Размер массива  $z$  должен быть выбран таким, чтобы массив  $a$  был занят не более чем на 90%.

Для выбора функции расстановки следует обратить внимание, что во многих практических применениях множество возможных ключей значительно больше числа допустимых значений индексов. В частности, обычно приходится исходить из того, что должна запоминаться только небольшая часть значений ключей, но при этом за-

ранее не известно, каковы эти значения. Закономерно, что в таких случаях пользователь заинтересован в том, чтобы функция расстановки по возможности равномерно отображала множество значений ключей на значения индексов.

Если сами ключи также заданы в виде натуральных чисел, например из интервала от 0 до  $s-1$ , где число  $s$  значительно больше числа  $z$ , то в качестве функции расстановки можно просто взять

$$h(i) = i \bmod z.$$

Эта функция фактически обладает таким свойством, благодаря которому значения ключей равномерно распределяются по области значений индекса. Отсюда следует, что с помощью трансформационного отображения возможно значения ключей однозначно отобразить на интервал натуральных чисел, а эту функцию можно использовать в качестве функции расстановки. Впрочем, если имеются какие-то дополнительные статистические данные о распределении ключей, отличном от равномерного, то следует использовать другую, более подходящую функцию расстановки. Следует также обратить внимание на то, что вычисление этой функции не должно быть слишком трудоемким.

Примером, когда невозможно исходить из равномерного распределения ключей, является запоминание в хэш-памяти слов из некоторого текста. При наивном подходе напрашивается следующий способ действий: последовательные буквы слова кодировать двоичными цифрами и слово текста хранить в полученной таким образом двоичной кодировке, а в качестве функции расстановки принять просто проекцию—например, в качестве значения функции принять код первой буквы слова. Однако этот способ, как правило, неудачен, так как он не сможет обеспечить равномерного распределения ключей по области значений индекса.

Для разрешения коллизий следует поступить следующим образом. Если при возникновении коллизии оба ключа должны быть запомнены в хэш-памяти, то дополнительно к собственно индексу для занесения в хэш-память должен быть найден заменяющий индекс. Здесь речь идет об *открытой адресации* в методе хэширования.

Предлагается и следующий, принципиально иной способ действий. На каждый индекс в хэш-памяти предусматривается занесение не одного содержимого, а целого их множества. Это может быть реализовано, например, путем образования списка из заносимых значений. Речь идет о *непосредственном сцеплении*. В этом случае после определения индекса для ключа надо просмотреть этот список и проверить, было ли занесение по этому индексу, и если да, то занос-

симый элемент должен быть внесен в этот список. Такой способ требует контроля за переполнением хэш-массива, а потому необходимости выделения дополнительной памяти для размещения приводящих к коллизии элементов, которые не удастся разместить непосредственно в хэш-памяти. В этом случае говорится о *закрытой адресации* в методе хэширования.

При открытой адресации дополнительные элементы при коллизии помещаются в самом хэш-массиве. Поиск места в хэш-массиве для занесения элемента в случае возникновения коллизии будем называть *зондированием*.

Если при вычислении значения индекса для заданного ключа выясняется, что по этому индексу уже занесен элемент данных с другим ключом, то по определенному правилу вычисляется следующий индекс, по которому и заносится элемент данных. При условии, что по этому индексу был уже занесен элемент данных, вычисляется следующий индекс и т.д.—до тех пор, пока не будет найдено свободное место в массиве.

При обращении к какому-либо элементу в хэш-массиве для его чтения может случиться так, что по индексу для заданного ключа находится элемент с другим ключом. В таком случае аналогично тому, как это делалось при занесении элементов данных, надо перебирать последовательность значений индекса, по которым мог быть записан этот элемент.

Различают *линейное* и *квадратичное* зондирование. При линейном зондировании позиции хэш-массива просматриваются с постоянным шагом, а при квадратичном, исходя из значения  $h(i)$ ,—значения индекса

$$h(i)+1 \bmod z, h(i)+4 \bmod z, h(i)+9 \bmod z, \dots, h(i)+j^2 \bmod z$$

Простой способ линейного зондирования для определения нового значения индекса в случае коллизии состоит в том, что значение индекса (по модулю  $z$ ) по мере необходимости увеличивается на 1, пока не будет найдено свободное место в массиве. Впрочем, этот способ имеет недостаток: при некоторой статистике скоплений могут возникнуть сгущения в хэш-массиве, что может привести к интенсивной загруженности области массива (вследствие чего потребуется длительный поиск свободного места для заносимого элемента), в то время как другие области будут заняты очень слабо.

Более подходящей будет такая функция разрешения коллизий, которая равномерно распределяет ключи по остальному множеству свободных мест. Однако этот способ может быть очень дорогим. Поэтому на практике ищут компромисс. берут, например, функцию,

которая, как это показано выше, распределяет ключи квадратичным образом. Впрочем, может случиться, что в процессе поиска по хэш-массиву не будет найдено свободного места для размещения элемента. При квадратичном зондировании будет просматриваться по меньшей мере половина массива, если в качестве его длины взято простое число.

Метод хэширования весьма эффективен в том случае, когда используются только рассмотренные операции над элементами множеств. Если же в определенных приложениях должны использоваться все элементы данных, например, обработка типа сортировки или такие операции над множествами, как объединение и пересечение, то этот метод менее удобен. Для того, чтобы и в этих случаях можно было эффективно работать с хэш-массивами, необходима трудоемкая предварительная сортировка.

Анализ статистики показывает, что метод хэширования чрезвычайно эффективен при небольшом заполнении хэш-массива. Даже его заполнение на 90% при достаточно удачно выбранном способе зондирования для занесения или выбора нужного элемента в среднем требует 2,56 шага зондирования. Это значение существенно зависит от степени загрузки массива. Поэтому размер хэш-массива следует выбирать таким, чтобы в процессе работы он заполнялся не более чем на 90%.

Отсюда вытекают и недостатки метода хэширования. Если размер массива выбрать слишком большим по отношению к числу фактически хранимых в нем элементов, то значительная часть выделенной памяти будет просто пропадать. Если же размер массива окажется слишком малым, то будет возникать слишком много коллизий, для их устранения придется просматривать длинные списки элементов и по затратам времени метод окажется неэффективным.

Таим образом, недостаток состоит прежде всего в том, что размер хэш-массива должен быть выбран и зафиксирован предварительно и этот размер не может быть динамически подогнан к числу фактически заносимых элементов.

Другой недостаток состоит в том, что весьма сложна процедура удаления элементов—особенно в тех случаях, когда используется техника размещения элементов, вызывающих коллизию. В этом случае при удалении элемента придется осуществлять перезапоминание элементов, которые размещались в памяти в результате разрешения коллизии, а это влечет за собой весьма сложные преобразования соответствующих списков.

Существует много различных вариантов метода хэширования.

Рафинированные способы хэширования получаются при так называемом “grid file”, когда в хэш-памяти размещается информация с двумерными и многомерными китчами с помощью функций хэширования. Для этого плоскость или пространство делят на растры точками, которые хранятся в таблице, и отсюда определяют функцию хэширования.

## **24. Информационная безопасность**

### **24.1. Геоинформатика и информационная безопасность**

Современная информация это ценный ресурс, а не просто отдельные данные. Современное информационное общество характеризуется тем, что информация стала объектом гражданско-правовых отношений вошла в рыночный оборот. На получение и систематизацию различных информационных продуктов и услуг, формирование информационных ресурсов - затрачиваются значительные финансовые средства. Это обуславливает рост ценности информации и необходимость организации и поддержки ее безопасности [63, 101, 115, 117, 118, 120, 129, 131].

Особенностью геоинформатики является сбор и интеграция широкого круга данных, применяемых в комплексной обработке для решения разных задач [58, 72, 73]. Это определяет преимущество ГИС, но с другой стороны создает условия для большего числа каналов доступа и информационных угроз в этой области.

Следует отметить, что информационная безопасность вообще и в геоинформатике в частности требует достаточно объемного изложения. Поэтому в данной главе дается краткое изложение этого вопроса. Более подробная информация приведена в работах [63, 101, 115, 117, 118, 120, 129, 131 и др.]

Термин “защита информации” в настоящее время отражает узкую часть проблем информационной безопасности.

В 60-е годы возникло понятие “компьютерной безопасности”, в 70-е - “безопасность данных”. В настоящее время более полным понятием, связанным с безопасностью в компьютерных системах и информационных технологиях, является понятие “информационная безопасность”.

Информатизация общества выдвигает требования по постепенному сокращению безвозмездного доступа к информационным ресурсам и обеспечение их защиты от несанкционированного распространения.

Необходимость защиты информации возрастает по мере увеличения ее ценности, с развитием открытых систем и технологий

телекоммуникаций, с возрастанием уровня информатизации общества и повышением технического уровня конечных пользователей компьютерных технологий.

**Информационной безопасностью** называют комплекс организационных, технических и технологических мер по защите информации от неавторизованного доступа, разрушения, модификации, раскрытия и задержек в доступе. Информационная безопасность, как правило, строится на системах информационной безопасности, которые включают комплекс методов по защите процессов обработки, хранения и представления информации.

Информационная безопасность дает гарантию того, что достигаются следующие цели:

- конфиденциальность информации;
- целостность информации и связанных с ней процессов;
- доступность информации, когда она нужна;
- учет всех процессов, связанных с информацией.

Важность необходимости организации информационной безопасности иллюстрирует следующий пример. Сотрудники Министерства обороны США решили проверить свою систему защиты на прочность. Считается, что информационная безопасность этого ведомства одна из лучших мире. Работу по взлому выполняли инженеры управления Defense Information System Agency (DISA). Им удалось вскрыть 7860 компьютерных систем (серверов и мэйнфреймов) из 8932 машин, составляющих компьютерный парк министерства. Характерно, что саму попытку несанкционированного доступа удалось зафиксировать лишь в 390 случаях.

Данный факт позволяет сделать вывод о том, что большинство организаций и компаний, уровень информационной безопасности которых значительно ниже, могут даже не подозревать о том, что кто-то успешно проник или пытался проникнуть в их компьютерные системы.

## 24.2. Система безопасности

Анализ мирового и отечественного опыта диктует необходимость создания целостной системы безопасности любой организации. Эта система должна, связывать оперативные, оперативно-технические и организационные меры защиты, использовать современные методы прогнозирования, анализа и моделирования ситуаций. Эта система безопасности должна включать и систему информационной безопасности [63, 115, 117].

Постоянно меняющаяся политическая, социальная и экономическая ситуация не позволяет заранее предусмотреть все возможные варианты реализации информационной безопасности и ограничиться некой совокупностью мер защиты.

Для учета динамики изменения внешней среды необходимо разрабатывать концепции информационной безопасности. Они должны быть построены на комплексе динамических моделей, представляющих собой систематизированное изложение целей, задач, принципов и способов достижения информационной безопасности.

Концепции информационной безопасности определяют стратегию, тактику информационной безопасности определяют цели предприятия в условиях внешней среды. [63]

Тактическая, оперативная деятельность и реализация информационной безопасности обусловлена реальными ресурсами предприятия (финансовыми, интеллектуальными, техническими).

При разработке систем защиты или безопасности используют следующую иерархию понятий по степени возрастания их надежности: технология информационной безопасности, система защиты, система информационной безопасности.

Для организации систем информационной безопасности необходимо классифицировать информационные объекты и субъекты доступа к ним. Множество информационных объектов можно условно разделить на две категории - с ограниченным доступом и общедоступные. Некоторые категории общедоступной информации могут со временем переходить в категорию с ограниченным доступом и наоборот.

Существуют категории общедоступной информации, доступ к которым не может быть ограничен никем ни при каких условиях, например сведения о состоянии окружающей среды.

### **24.3. Показатели защищенности средств вычислительной техники от несанкционированного доступа**

Современные геоинформационные и телекоммуникационные технологии базируются на использовании средств вычислительной техники (СВТ). Защищенность СВТ во многом определяет защищенность информационных технологий. Это является основанием для определения оценки защищенности СВТ.

Классификацию средств вычислительной техники по уровням защищенности от несанкционированного доступа (НСД) устанавливает руководящий документ, выпущенный Гостехкомиссией [33], таблица 24.1. Он определяет перечень показателей защищенности и совокупность описывающих их требований.

Таблица 24. 1. Распределение показателей защищенности по классам СВТ

Наименование показателя	Класс защищенности					
	6	5	4	3	2	1
Дискреционный принцип контроля доступа	+	+	+	=	+	=
Мандатный принцип контроля доступа	-	-	+	=	=	=
Очистка памяти	-	+	+	+	=	=
Изоляция модулей	-	-	+	=	+	=
Маркировка документов	-	-	+	=	=	=
Защита ввода и вывода на отчуждённый физический носитель	-	-	+	=	=	=
Сопоставление пользователя с устройством	-	-	+	=	=	=
Идентификация и аутентификация	+	=	+	=	=	=
Гарантии проектирования	-	+	+	+	+	+
Регистрация	-	+	+	+	=	=
Взаимодействие пользователя с КСЗ	-	-	-	+	=	=
Надёжное восстановление	-	-	-	+	=	=
Целостность КСЗ	-	+	+	+	=	=
Контроль модификации	-	-	-	-	+	=
Контроль дистрибуции	-	-	-	-	+	=
Гарантии архитектуры	-	-	-	-	-	+
Тестирование	+	+	+	+	+	=
Руководство пользователя	+	=	=	=	=	=
Руководство по КСЗ	+	+	=	+	+	=
Текстовая документация	+	+	+	+	+	=
Конструкторская (проектная) документация	+	+	+	+	+	+



Обозначения:

“-” - нет требований к данному классу

“+” - новые дополнительные требования

“=” - требования совпадают с требованиями к СВТ предыдущего класса

“КСЗ” - комплекс средств защиты

Установлено семь классов защищенности СВТ от НСД к информации. Самые низкие требования предъявляются к системам, соответствующим седьмому классу, самые высокие — к первому.

Показатели защищенности и установленные требования к классам приведены в табл.24.1.

Наборы требований, приведенные в таблице 24.1. к показателям каждого класса являются минимально необходимыми. В таблице приведены показатели шести классов. Седьмой класс присваивают СВТ, к которым предъявлялись требования по защите от НСД к информации, но при оценке защищенность СВТ оказалась ниже уровня требований шестого класса.

#### **24.4. Классы защищенности автоматизированных систем**

Автоматизированная система, например геоинформационная, является более сложной системой по сравнению со средствами вычислительной техники. Она является человеко-машинной системой с множеством различных каналов доступа возможной утечки информации. Документы Гостехкомиссии устанавливают девять классов защищенности автоматизированных систем (АС) от НСД. Каждый класс характеризуется определенной совокупностью требований к средствам защиты

Классы подразделяются на три группы, отличающиеся спецификой обработки информации в автоматизированных системах.

Группы определяются на основании следующих признаков:

- наличие в АС информации различного уровня конфиденциальности;
- уровень полномочий пользователей АС на доступ к конфиденциальной информации;
- режим обработки данных в АС (коллективный или индивидуальный).

В пределах каждой группы соблюдается иерархия классов защищенности АС.

Таблица 25.1. Классы защищенности АС

Подсистемы и требования	Классы								
	3Б	3А	2Б	2А	1Д	1Г	1В	1Б	1А
	Гр.3		Гр.2		Группа 1				
1. Подсистема управления доступом									
1.1. Идентификация. Проверка и контроль доступа субъектов в систему	+	+	+	+	+	+	+	+	+
К терминалам, ЭВМ, узлам сети ЭВМ, каналам связи, внешним устройствам ЭВМ,				+		+	+	+	+
К программам				+		+	+	+	+
К томам, каталогам, файлам, записям, полям записей				+		+	+	+	+
1.2. Управление потоками информации				+			+	+	+
2. Подсистема регистрации и учёта									
2.1. Регистрация и учёт: входа/выхода субъектов доступа в/из системы (узла сети)	+	+	+	+	+	+	+	+	+
Выдачи печатных (графических) выходных документов		+		+		+	+	+	+
Запуска/завершения программ и процессов (заданий, задач).				+		+	+	+	+

Доступа программ субъектов к защищаемым файлам, включая их создание и удаление, передача по линиям и каналам связи,				+		+	+	+	+
Доступа программ субъектов, доступа к терминалам ЭВМ, узлам сети ЭВМ, каналам связи, внешним устройствам ЭВМ, программам, каталогам, файлам, записям, полям записей,				+		+	+	+	+
Изменение полномочий субъекта доступа,							+	+	+
Создаваемых защищаемых объектах доступа.				+			+	+	+
2.2. Учёт носителей информации	+	+	+	+	+	+	+	+	+
2.3. очистка (обнуление, обезличивание) освобождаемых областей оперативной памяти ЭВМ и внешних накопителей.		+		+		+	+	+	+
2.4. Сигнализация попыток нарушения защиты							+	+	+
3. Криптографическая подсистема									
3.1. Шифрование конфиденциальной информации				+				+	+

3.2. Шифрование информации, принадлежащей различным субъектам доступа (группам субъектов) на разных ключах									+
3.3. Использование аттестованных (сертифицированных) криптографических средств				+					+
4. Подсистема обеспечения целостности.									
4.1. Обеспечение целостности программных средств и обрабатываемой информации	+	+	+	+	+	+	+	+	+
4.2. Физическая охрана средств вычислительной техники и носителей информации	+	+	+	+	+	+	+	+	+
4.3. Наличие администратора (службы) защиты информации в АС				+				+	+
4.4. Периодическое тестирование СЗИ НДС	+	+	+	+	+	+	+	+	+
4.5. Наличие средств восстановления СЗИ НДС	+	+	+	+	+	+	+	+	+
4.6. Использование сертифицированных средств защиты		+		+				+	+

Обозначения:

“+” – требование к данному классу присутствует

“СЗИ НДС” – система защиты информации от несанкционированного доступа.

Класс, соответствующий высшей степени защищенности обозначается индексом «А», более низкий класс обозначается индексом «Б» и т.д.

Обозначение группы и класса защищенности записывается как *NA.*, где *N* — номер группы (от 1 до 3).

Первая группа включает многопользовательские автоматизированные системы, в которых допускается одновременная обработка и/или хранение информации разных уровней конфиденциальности.

Пользователи этой группы имеют разные права доступа. Группа содержит пять классов — 1Д, 1Г, 1В, 1Б и 1А.

Вторая группа включает автоматизированные системы, в которых пользователи имеют одинаковые полномочия доступа ко всей информации, обрабатываемой и/или хранимой в АС на носителях различного уровня конфиденциальности. Группа содержит два класса — 2Б и 2А.

Третья группа включает системы, в которых работает один пользователь, допущенный ко всей информации, размещенной на носителях одного уровня конфиденциальности. Группа содержит два класса — 3Б и 3А. В табл. 24.2 приведены требования к подсистемам защиты для каждого класса.

## 24.5. Каналы доступа к информационным системам

Анализ возможных каналов НДС к информационным системам показывает, что данные каналы необходимо разделить на два вида: контролируемые и неконтролируемые.

К контролируемым каналам доступа информационной системы можно отнести:

- · терминалы пользователей;
- · средства отображения и документирования информации;
- · средства загрузки программного обеспечения;
- · технологические пульты и органы управления;
- · внутренний монтаж аппаратуры;
- · побочное электромагнитное излучение;

К неконтролируемым каналам НДС к информации информационной системы следует отнести:

- · машинные носители программного обеспечения и информации выносимые за пределы информационной системы;

- · долговременные запоминающие устройства с остатками информации, выносимыми за пределы информационной системы;
- · внешние каналы связи;
- · мусорная корзина.

Кроме того, учитывая множество пользователей, допущенных к терминалам и информации внутри информационной системы, требуется создание встроенной системы контроля и разграничение доступа.

Разграничение доступа пользователей к информации информационной системы должно производиться в соответствии с регламентом, установленным администрацией.

На контролируемых каналах доступа все цепи и тракты контроля аппаратно, программно и организационно должны сходиться на одном рабочем месте службы безопасности информации или администратора безопасности информационной системы.

#### **24.6. Группы методов несанкционированного доступа**

При организации систем безопасности необходимо принять во внимание основные группы методов, используемые для совершения несанкционированного доступа. Перечислим эти группы:

**1. Имитация** - группа методов основана на подмене информации в процессе ее ввода или вывода. Например, при вводе документы могут быть заменены фальшивыми, вместо рабочих дискет подсунуты чужие, и данные могут быть сфальсифицированы.

**2. Сканирование** - анализ всей доступной информации. Эти способы включают просмотр мусорных корзин (включая компьютерные); анализ оставленных на столе дискет, CD, DVD; анализ "стертых" файлов и т.п.

**3. Троянский конь** - группа методов, основанных на подмене известной программы или информационного продукта аналогом, содержащим дополнительные функции. Программа, выполняющая известные полезные функции, пишется таким образом, что содержит дополнительные скрытые функции, которые можно использовать независимо от официального владельца программы.

**4. Люк** - группа методов (упрощенный вариант предшествующей группы), основанных на использовании скрытого программного или аппаратного механизма, позволяющего обойти методы защиты в системе и войти в системы. Этот механизм активируется некоторым неочевидным образом. Иногда программа пишется таким образом, что специфическое событие, например, число транзакций, обрабо-

танных в определенный день, вызовет запуск неавторизованного механизма..

5. **Технология салямн** - группа методов, основанных на НСД совершаемых небольшими частями, настолько маленькими, что они незаметны. Обычно эта технология сопровождается изменением компьютерной программы. Например, платежи могут округляться до нескольких центов, и разница между реальной и округленной суммой поступать на специально открытый счет злоумышленника.

6. **Суперотключение** - специфическая группа методов, основанных на вход в систему в сбойных ситуациях. Названа по имени программы, использовавшейся в ряде компьютерных центров, обходившей системные меры защиты и использовавшейся при аварийных ситуациях. Владение этим методом дает возможность в любое время получить доступ к компьютеру и информации, находящейся в нем.

Помимо анализа возможных методов НСД при создании систем информационной безопасности необходим индивидуальный анализ информационной системы с целью выявления ее уязвимых мест.

## 24.7. Информационные угрозы

Основным принципом создания системы информационной безопасности должно стать обеспечение заданного уровня защищенности от возможных угроз при минимальной стоимости средств и систем защиты.

Анализ угроз проведенных агентством национальной ассоциацией информационной безопасности (National Computer Security Association) в 1998 г. в США выявил следующую статистику основных угроз:

сбои оборудования -	43%;
неумелые или неправильные действия персонала	26%;
вредительские действия собственных сотрудников	15%;
внешние атаки по сети (Интернет)	8%;
воздействие компьютерных вирусов	5 %.

Эти угрозы могут определять действия и мероприятия по созданию информационной безопасности систем, включая геоинформационные.

## 24.8. Предотвращение угроз

Идентификация, классификация и анализ возможных угроз повышает эффективность защиты. Как правило, ожидаемая угроза распознается быстрее, чем неожиданная.

Своевременное обнаружение угрозы НСД ускоряется сле-

ствующими мерами:

- ранжированием угроз по вероятности осуществления;
- составлением перечня субъектов (лиц и организаций) – потенциальных посягателей;
- мониторингом состояния наиболее уязвимых направлений деятельности субъекта или объекта;
- применением технических средств обнаружения (ТСО) угрозы посягательства с характеристиками, позволяющими минимизировать (в пределах заданной стоимости ТСО) время обнаружения НСД;
- оптимальным сочетанием применения физической охраны и ТСО.

#### **24.9. Политика безопасности и уровни ее реализации**

Система информационной безопасности должна постоянно совершенствоваться. Реализация концепций безопасности осуществляется посредством политики безопасности.

**Политикой безопасности** называют комплекс мер и активных действий по управлению и совершенствованию систем и технологий безопасности, включая информационную безопасность.

В настоящее время политика безопасности во многих странах осуществляется на четырех уровнях: законодательном, административном, процедурном, программно-технологическом.

**Законодательный уровень** политики безопасности определяется нормативно-правовой информационной базой страны. Он является важнейшим для обеспечения информационной безопасности.

На этом уровне выделяют две группы мер:

- меры, направленные на создание и поддержание неаппетитного (в том числе карательного) отношения к нарушителям информационной безопасности;
- меры, способствующие повышению образованности общества в области информационной безопасности, помогающие в разработке и распространении средств обеспечения информационной безопасности.

**Законодательный уровень** определяет стратегию государства в области информационной безопасности, а также ту меру внимания и количество ресурсов, которую правительство данной страны считает целесообразным выделить на информационную безопасность.

**Административный уровень** политики информационной безопасности определяется совокупностью организационно-управленческих мероприятий предпринимаемых на каждом пред-



приятии или фирме. Он обеспечивается совокупность документированных управленческих решений, направленных на защиту информационных и ассоциированных с ними ресурсов.

Административный уровень задает стратегию организации в области информационной безопасности, а также ту меру внимания и количество ресурсов, которую руководство предприятия считает целесообразным выделить. Он соответствует высшему уровню управления предприятием.

Политика безопасности строится на основе анализа рисков, которые признаются реальными для информационной системы организации. Когда риски проанализированы и стратегия защиты определена, составляется программа, реализация которой должна обеспечить информационную безопасность. Под эту программу выделяются ресурсы, назначаются ответственные, определяется порядок контроля выполнения программы и т.п.

Административный уровень информационной безопасности наименее развит в нашей стране. Нет законов, обязывающих организации иметь политику безопасности. Мало кто из руководителей знает, что такое политика безопасности, еще меньшее число организаций такую политику реализуют.

**Процедурный уровень** включает меры безопасности, реализуемые на низшем (операционном) уровне управления. Несмотря на то, что в отечественных организациях накоплен богатый опыт составления и реализации процедурных (регламентных) мер, они нуждаются в существенном пересмотре, поскольку возникли задолго до появления компьютерных информационных технологий.

Выделяют следующие группы процедурных мер, направленных на обеспечение информационной безопасности:

- управление персоналом;
- физическая защита;
- поддержание работоспособности;
- реагирование на нарушения режима безопасности;
- планирование восстановительных работ.

**Программно-технический уровень** включает меры безопасности, реализуемые за счет программно технологических средств. Собственно это уровень реализации, в то время как остальные уровни - уровни организации..

В рамках этого уровня должны быть доступны по крайней мере следующие механизмы безопасности:

- идентификация и проверка подлинности (аутентификация) пользователей;

- управление доступом;
- протоколирование и аудит;
- криптография;
- (межсетевое) экранирование.

#### **24.10. Методы защиты от компьютерных вирусов**

Программно-аппаратный сторож является одним из методов защиты. Он включает устройство и комплекс программ, предназначенные для защиты данных и системных областей дисков от любого несанкционированного изменения. Устройство обычно реализуется в виде дополнительной платы, подключаемой к шине компьютера.

Одним из основных методов борьбы с вирусами является организация порядка на рабочих местах, включающая регулярную профилактику. Антивирусная профилактика предполагает соблюдение некоторых правил, позволяющих значительно снизить вероятность заражения вирусом и потери каких – либо данных.

1. Компьютер должен содержать антивирусную программу с достаточно большой и актуальной на момент установки антивирусной базой.
2. Антивирус должен работать в режиме on-line или «на лету», что позволит автоматически детектировать все файлы вставляемых дискет и фиксировать вирусоподобные изменения программ при отсутствии действий по проверке со стороны пользователя.
3. Поскольку антивирус сам является потенциальным разносчиком вируса, целесообразно иметь антивирусные программы на CD. Это обусловлено тем, что в отличие от магнитных дисков, вирус, при всем своем желании, на оптический диск себя не запишет.
4. При специальной проверке больших дисков и вероятности заражения антивирусной программы целесообразно использование связки двух программ: первая проверяет диски; вторая детектирует первую.
5. Необходима регулярное обновление антивирусной базы, не реже одного раза в месяц.
6. В режиме MS DOS необходимо регулярная проверка содержимого "мусорной корзины" на предмет появления новых "нестирающихся" файлов.
7. Необходимо ограничить число директорий общего доступа при работе с сетью.

## 24.11. Виды возможных последствий НСД к информации ГИС

Виды возможных последствий НСД к информации ГИС следующие:

**Утрата** заключается в стирании, искажении, уничтожении или хищении информации, находящейся в процессе обработки или хранения в ГИС. Опасность ее заключается в безвозвратной потере ценной информации.

**Модификация** включает модификацию данных и модификацию программ.

**Модификация данных** заключается в изменении информации на ложную, которая корректна по форме и содержанию, но имеет другой смысл. Навязывание ложной информации при отсутствии законной и недоведение законной информации до получателя можно также считать ее модификацией. Примером модификации может служить изменение даты позиционное смещение системы координат и т. д. Опасность модификации заключается в возможности организации утечки секретной или/и передачи ложной информации в дальнейшую обработку и использование.

**Модификация программ** заключается в изменении свойств программ при внешней корректности их работы. Например, это может быть заражение вирусом, повышение вычислительной погрешности или внесение в данные в ходе обработки систематических или случайных погрешностей.

**Утечка информации** заключается в несанкционированном ознакомлении постороннего лица с секретной информацией.

**Искажение** - результат обработки данных не сертифицированными программами, которые вносят систематические погрешности в позиционные данные.

**Взлом системы** - модификация программного обеспечения, обеспечивающая его работу без ключей защиты, что дает возможность простого анализа данных ГИС, имеющих специальные форматы, характерные только для данной ГИС и нечитаемые другими программными средствами.

Чем большее количество возможных каналов несанкционированного доступа перекрыто средствами защиты, тем выше уровень безопасности информации, обрабатываемой в данной ГИС и тем меньше вероятность преодоления защиты потенциальным нарушителем. Очевидно, что в реальной геоинформационной технологии структура защиты должна носить многозвенный и многоуровневый характер.

Количество перекрываемых каналов доступа при этом будет зависеть от заданной квалификации нарушителя.

## **24.12. Функции системы информационной безопасности в геоинформатике**

Целью функционирования системы информационной безопасности в геоинформатике является обеспечение безопасности геоинформационной системы, защита точности и целостности информации, минимизация последствий НСД. Информационная безопасность требует учета всех событий, в ходе которых информация создается, модифицируется, к ней обеспечивается доступ или она распространяется [63, 78, 117, 120, 129].

На процедурном уровне можно определить четыре функции системы информационной безопасности ГИС.

**Предотвращение** - только авторизованный персонал имеет доступ к информации и ГИС-технологии.

**Обнаружение** - обеспечивается раннее обнаружение преступлений и злоупотреблений, даже если механизмы защиты были обойдены.

**Ограничение** - уменьшается размер последствий НСД, если оно все-таки произошло.

**Восстановление** - обеспечивается эффективное восстановление информации при наличии документированных и проверенных планов по восстановлению

Безопасность геоинформационных технологий — это гораздо более широкое и глубокое понятие, чем защита информации. Угрозу представляет не только возможность несанкционированного доступа к ресурсам ГИС, но и возможность через информационные технологии нанесения ущерба тем, в интересах кого они применяются. Эти два аспекта определяют внутреннюю и внешнюю безопасность ГИС-технологий.

С учетом изложенного, под безопасностью ГИС-технологий предлагается понимать их способность обеспечивать защиту информационных ресурсов и технологий от действия внешних и внутренних, случайных и преднамеренных угроз, а также выполнять предписанные им функции без нанесения ущерба потребителям информации.

Поскольку состояние внешней среды по отношению к ГИС-технологиям меняется, необходимо периодически проводить их внешний и внутренний аудит. Внутренний аудит связан с анализом системных журналов, внешний с оценкой безопасности ГИС и ГИС-технологий к внешним угрозам.

Оценка безопасности ГИС-технологий производится с целью проверки соответствия достигнутого уровня безопасности заданному в ТЗ на разработку ГИС-технологий, а также соответствия требованиям стандартов и нормативных документов в отношении ГИС-технологий данного класса.

При определении оценки безопасности используются спецификации, нормативные материалы, программно-методический аппарат проведения оценок и принятые стандарты и критерии.

В настоящее время общеупотребительным подходом к построению критериев оценки безопасности ГИС-технологий является использование совокупности определенным образом упорядоченных качественных требований к функциональным механизмам обеспечения безопасности, их эффективности и гарантированности реализации. Такая совокупность гармонизированных стандартов и критериев оценки называется профилем.

Результирующая оценка уровня безопасности ГИС-технологий имеет качественное выражение. Такой подход положен в основу всех реально действующих нормативных документов по оценке безопасности. Пока ни в одном из нормативных документов по оценке безопасности количественные показатели не применяются, хотя в некоторых допускается их использование.

Получение количественных характеристик безопасности было бы, несомненно, весьма полезным, особенно для сравнения различных проектов обеспечения безопасности, анализа влияния угроз и отдельных механизмов на общий уровень безопасности, учета изменения безопасности в процессе жизненного цикла ГИС-технологий. Однако для того, чтобы количественный показатель корректно использовать, он должен иметь объективную интерпретацию, однозначную зависимость от отдельных аспектов безопасности.

Специфика информационных технологий как объекта оценки определяется присутствием в них в качестве определяющей компоненты программного обеспечения. Ввиду большой функциональной, структурной и логической сложности программного обеспечения на практике невозможно в полном объеме оценить его поведение во всем возможном диапазоне его применения.

Предметная область безопасности в геоинформатике обладает той особенностью, что при ее оценке приходится применять как объективные, так и субъективные критерии. Оцениваемые характеристики безопасности ГИС-технологий могут иметь как детерминированную, так и случайную природу. Некоторые исследуемые элементы ГИС-технологий, как, например, преднамеренные закладки, име-

ют уникальное представление и скрытый характер, что также затрудняет проведение оценок безопасности ГИС-технологий.

### **24.13. Оценки характеристики информационной безопасности ГИС**

В силу указанных обстоятельств получение интегральных количественных оценок безопасности ГИС-технологий является проблематичным. Это не исключает использования отдельных количественных показателей там, где это целесообразно. Например, эти показатели могут использоваться для оценки уровня механизмов криптографической и парольной защиты, времени модельного эксперимента, времени обнаружения НСД и др.

В качестве общих требований к нормативно-методической базе оценки безопасности ГИС-технологий можно выделить следующие:

- универсальность, способность обеспечивать оценку безопасности любых видов ГИС-технологий и отдельных их компонентов;
- гибкость, способность формирования требований и получения оценок безопасности ГИС-технологий, максимально учитывающих особенности их применения;
- конструктивность, способность объективным образом оценивать уровень безопасности ГИС-технологий и влиять на процесс ее обеспечения;
- приемственность, способность интерпретировать результаты оценок, полученных в других системах оценки безопасности ГИС-технологий;
- расширяемость, способность наращивания системы критериев и показателей без нарушения их общего построения.

Рассмотрение существующих систем нормативных документов в области сертификации ГИС-технологий показывает, что в своей основе они в значительной мере этим требованиям не удовлетворяют, поэтому необходимо их совершенствование.

Применение средств защиты на уровне приложения ГИС-технологий наиболее надежной, но в тоже время наиболее сложный вариант, так как реализовать его может только сам разработчик приложения. Этот метод дает более строгий контроль доступа к данным и процессам обработки, поскольку позволяет заложить в систему алгоритмы всевозможных проверок, отличные от стандартных, хорошо известных хакерам. Кроме того, такой подход позволяет создать систему самотестирования - метода анализа программного продукта им самим, реагирующего на все попытки проникновения

извне.

Некоторые разработчики коммерческих приложений вместе с основными продуктами поставляют собственные программы администрирования своих же приложений (прикладные утилиты администрирования), что позволяет достичь более высокой степени защиты и контроля работающих приложений. При этом администратору системы не нужны внешние программы администрирования, так как вся необходимая информация (списки пользователей, их пароли и права доступа) контролируется упомянутыми утилитами. К сожалению, в области ГИС-технологий такие работы ведутся мало.

Надежная система безопасности должна не только выполнять оперативные функции предотвращения НСД, но и функции аудита, фиксируя в аудиторском журнале все подозрительные события за время работы системы. Кроме того, что особенно важно для ГИС-технологий, система информационной безопасности должна включать периодическое резервное копирование. В ГИС обрабатывается много разнообразных взаимосвязанных данных (изображения, таблицы, графика), поэтому и повреждение какого либо из этой совокупности или нарушение связей приведут к невозможности дальнейшей работы.

Таким образом проблема организации информационной безопасности стоит особенно остро в ГИС -технологиях, если учесть, что они планируются к использованию и уже применяются как системы регионального и муниципального управления.

#### **24.14. Защита информационных ресурсов в геоинформатике**

Информационная безопасность в геоинформатике и ГИС должна быть организована комплексно, а не односторонне типа автономных “защита данных”, “сетевая защита”, “применение паролей”, “разграничение доступа” и т.п.

Появление ГИС как интегрированных систем с интеграцией технологий и интеграцией данных поставило еще одну проблему - проблему защиты качества информации и качества программ, обрабатывающих эту информацию.

Таким образом, защита информации в геоинформатике включает:

- защиту данных;
- защиту программ;
- защиту качества информации.

Наряду с внешними факторами, требующими защиты, появи-

пись внутреннее.

Для построения целостной системы защиты информации, предотвращения несанкционированного доступа к данным и программам в любой среде необходимо определить ее особенности и вытекающую из этих особенностей специфику в организации защиты информации.

Анализ геоинформационных технологий позволяет выделить специфичность организации данных и технологий и возможные каналы доступа к ГИС или ее информации.

**Информационная безопасность** ГИС - технологий определяется единой системой взаимосвязанных мер, обеспечивающих перекрытие возможных каналов несанкционированного доступа (НСД) от воздействий, направленных на утрату, модификацию, искажение, взлом системы и утечку информации.

Наступление одного из этих событий, не предусмотренных штатным режимом работы ГИС, рассматривается как факт совершения НСД.

#### **24.15. Уровни информационной защиты в ГИС**

Информационная безопасность тем выше, чем большее число дополняющих друг друга уровней защиты она включает. Каждый уровень защиты выполняет определенные функции. При этом одни уровни обеспечивают отражение внешних угроз, другие исключают внутренние причины порчи информации.

В геоинформатике информационная защита должна обеспечиваться на следующих уровнях:

- **концептуальном** - на уровне концепции разработки ГИС или геоинформационной технологии должны быть заложены принципы информационной безопасности до составления проекта и реализации программного обеспечения;
- **структурном** - на уровне построения структуры системы или технологии должна быть обеспечена внутренняя непротиворечивость методов анализа и обработки (отражение внутренних угроз) и обеспечена возможность контроля внешнего НСД (отражение внешних угроз);
- **алгоритмическом** - для повышения надежности сохранения результатов обработки результаты предыдущего шага обработки данных должны сохраняться до выполнения текущего шага (система отката);
- **программном** - качество обработки данных проверяется системой внутренних и внешних тестов;



- **программном** - пользователь не может самостоятельно запустить или иметь доступ к программному средству без санкции менеджера программ;
- **программном**- пользователь не имеет возможности самостоятельно заменить или модифицировать программное средство, а используемое обладает средствами защиты;
- **программном** - специальное системное программное обеспечение должно контролировать состояние векторов прерываний (особенно в многопользовательской системе), портов и т.д. с целью исключения доступа к процессу обработки или к данным со стороны программ-шпионов;
- **аппаратном** - система ключей, исключающих доступ к программному обеспечению;
- **физическом** - пользователь не имеет доступа к носителям информации.

#### **24.16. Тактика и стратегия защиты информации в геоинформатике**

Основные тактика и стратегия защиты информации от НСД в геоинформатике заключаются в выполнении следующих задач:

- · предупреждении и контроле попыток НСД;
- · своевременном обнаружении, определении места и блокировки не санкционированных действий;
- · регистрации и документировании события;
- · установлении и устранении причины НСД;
- · ведении статистики и прогнозировании НСД.

Как уже говорилось, информация в ГИС носит временной характер. Поэтому она постоянно должна обновляться.

При обновлении, в частности, с помощью передач данных по сети возникают дополнительные угрозы безопасности информации.

На каналах связи (в отличие от элементов ГИС) нарушитель ничем не рискует, особенно при пассивном перехвате информации, поэтому прочность защиты здесь должна быть особенно высокой. От активного вмешательства нарушителя в процесс обмена информацией между элементами ГИС и сети нужно применять систему обнаружения и блокировки НСД. Но и при этом риск нарушителя по-прежнему невелик, так как у него и в этом случае (по причине сложности определения его места пребывания) остается достаточно времени на то, чтобы отключиться и уничтожить свои следы.

## 24.17. Информационная безопасность и Интернет

Информационная безопасность и Internet - несовместимы по самой природе Internet. Она родилась как корпоративная сеть, а в настоящее время объединяет корпоративные и ведомственные сети (государственные, военные и т.д.) и рядовых пользователей. Это упрощает доступ в сеть для любого пользователя. Чем проще доступ в сеть, тем слабее ее информационная безопасность. Это создает такие условия, что пользователь может даже и не узнать о копировании у него файлов и программ, не говоря уже о возможности их порчи и корректировки.

Повышенный интерес к Интернет обусловлен ее доступностью и коммерциализацией информационных ресурсов.

Если защиту компьютеров можно обеспечить с помощью шлюзов, то передаваемые данные фактически открыты. Как только организации подключаются к сети, они могут быть атакованы злоумышленниками.

Как правило, посторонние лица предварительно проверяют сети организаций на возможность проникновения в них путем методического сканирования систем в них на наличие уязвимых мест. Для решения этих задач используют средства автоматического зондирования, то есть программы, которые сканируют все хосты, присоединенные к сети организации. Эта деятельность называется иногда *зондированием сети*.

Зондирование сети даже, если не приводит к большому финансовому ущербу, может нанести вред репутации атакованной организации.

Сети организаций, которые игнорируют эти проблемы, подвергают себя риску утечки или порчи ценной информации. Организации, в которых информационной безопасности уделяется недостаточное внимание, могут подвергаться атакам при наличии или выявлении новых уязвимых мест в системе защиты. Таким образом, в Интернет в *явной форме присутствуют информационные угрозы безопасности*.

## 24.18. Условия безопасности передачи данных по сети

Практически каналы связи в сети защитить не возможно, поэтому целесообразно строить защиту информации и сопровождающих ее служебных признаков на основе специальных криптографических преобразований. Одним из способов передачи геоинформационных данных по сетям может быть кодограмма.

При этом после соединения между двумя абонентами геоин-

формационной сети необходимо обеспечить следующие условия:

- а) получатель сообщения должен быть уверен в истинности источника данных;
- б) получатель должен быть уверен в истинности полученных данных;
- в) отправитель должен быть уверен в доставке данных получателю;
- г) отправитель должен быть уверен в истинности доставленных получателю данных.
- д) отправитель и получатель должны быть уверены в том, что с доставленной информацией в сообщении никто кроме них не ознакомился;
- е) отправитель и получатель должны быть уверены, что никому, кроме них и специального посредника, факт передачи сообщения между ними не известен;
- ж) получатель должен быть уверен, что отправитель - это то лицо, за которое себя выдает;
- з) отправитель должен быть уверен, что получатель - то лицо, которое ему необходимо для передачи сообщения.

Для обеспечения возможности контроля и разграничения доступа целесообразно для всех участников обмена информацией, помимо условных номеров, присвоить переменные идентификаторы в виде паролей, передаваемых в открытом виде и подлинность которых будет обеспечиваться механизмом цифровой подписи. Тем абонентам, которым присвоены соответствующие полномочия, должны быть предоставлены соответствующие значения паролей и закрытых ключей шифрования

Все вышесказанное позволяет предложить следующую группу средств для обеспечения безопасности геоинформационных данных, передаваемым по каналам связи, ориентированных на выполнение приведенных выше условий:

- а) средства формирования цифровой подписи сообщений;
- б) средства шифрования передаваемых данных;
- в) средства обеспечения цифровой подписи служебных признаков передаваемой информации, включая адреса и маршруты сообщения, а также получение отправителем и посредником квитанции от получателя.

Организационные мероприятия по защите информации должны предусматривать разграничение полномочий пользователей и включать средства регистрации доступа к информации в ГИС.

## **25. Стандартизация информационных продуктов и систем в геоинформатике**

Специалист в области геоинформатики, разработки или применения информационных систем и технологий любой специальности должен обладать знанием методов стандартизации и разработки программного обеспечения, оценки его надежности и правил его проверки.

### **25.1. Роль стандартизации в сфере геоинформатики**

Одной из причин неэффективного использования современных средств автоматизации является построение своих собственных информационных моделей с нарушением общепринятых стандартов. В результате такого подхода применение подобных информационных моделей при решении в комплексных задач оказывается малоэффективным.

Стандарты и стандартизация выполняют много важных функций в современном информационном обществе и в информационных технологиях и системах особенно. Они необходимы для эффективного обмена информацией между информационными или геоинформационными системами и обеспечения качества информационных продуктов [96, 97, 107, 127].

Это определяет необходимость изучения и применения методов стандартизации информационных средств для повышения эффективности геоинформационных технологий.

В настоящее время доминирующую роль для рынка информационных средств, информационных продуктов и для многих других сфер бизнеса и производства играют новые информационные технологии (НИТ), называемые также как информационные и коммуникационные технологии (ИКТ) (Information and Communication Technology (ICT)).

Современное динамическое развитие компьютерных технологий требует постоянного введения новых информационных стандартов и обновления существующих.

Современное общество - это информационное общество, которое развивается тем успешней, чем выше в нем уровень информатизации. Стандарты - определяющая предпосылка для развития информационного общества. Без них, соответствующая интеграция общества, способность к взаимодействию систем и услуг не будут возможны в будущем. При этом следует иметь ввиду, что в настоящее время концепция и значение стандарта изменяется.

## 25.2. Два типа стандартов в области новых информационных технологий

Динамика развития НИТ приводит к необходимости постоянного обновления и введения новых стандартов. В противном случае происходит появление “двойных стандартов”, одни из которых признаны соответствующими органами стандартизации и существуют формально, другие действуют фактически, будучи непризнанными как официальные документы [127].

Обычно де-юре ссылаются на стандартные рекомендации, выпущенные международными органами стандартизации типа ITU-T и ISO и др.; де-факто в качестве стандартов используют рекомендации или технические спецификации, принятые рынком и пользователем на международном уровне (даже в тех случаях, когда они выпущены одиночной компанией или группой компаний).

Операционная система Windows или Netscape браузеры - типичные примеры международных фактических стандартов, не утвержденных официально, но признанных специалистами и рынком.

Основные проблемы, с которыми сталкиваются де-юре стандарты определяются длинными задержкам для их окончательного формального утверждения. При этом довольно часто чрезмерно академические и теоретические сопровождаемые подходы фактически не поддерживаются изготовителями информационных средств и продуктов.

Примером официальных стандартов не получивших широкой поддержки на практике могут служить OSI (Open System Interconnection) (стандарты открытых систем связи). Множество протоколов, которые легли в основу этих стандартов, несмотря на всесторонне обоснованные концепции не привели к практической их реализации.

В нашей стране примером официального, но не получившего признания стандарта может служить ГОСТ 19.008-85 “P-СХЕМЫ АЛГОРИТМОВ И ПРОГРАММ”. Стандарт определяет условные графические обозначения алгоритмов и правила их выполнения. Из-за сложности этот стандарт практически не применялся. Вместо него применялись и применяются до настоящего времени более простые стандарты, входящие в ЕСТПП, выпущенные в 1974. Они соответствуют стандартным обозначениям, применяемым в учебных курсах по информатике или по программированию. Этот стандарт отвечает требованиям структурного анализа и программирования, чем объясняется их распространение и долголетие.

### 25.3. Концепции разработки стандартов

Для динамического развития информационных средств и технологий в геоинформатике необходим учет главных тенденций разработки информационных и программных средств.

Одной из главных характеристик любой продукции, включая информационную, является жизненный цикл [103, 107].

**Стандарт** - показатель жизненного цикла технологии, и эта концепция должна быть заложена при его разработке - таково мнение европейских специалистов по стандартизации в области НИТ [127].

В этом случае различие между де-юре и де-факто стандартами менее значительно для пользователей, но в любом случае стандарты с одной стороны не должны являться негативным средством влияния на введение инноваций, с другой - должны ускорять и способствовать рыночным отношениям.

В процессе стандартизации решаются следующие задачи: определение объекта стандартизации, определение области стандартизации для данного объекта, определение основных целей стандарта, определение терминологических категорий стандарта, соответствие стандарта международным аналогам или требованиям в данной области стандартизации .

Существуют два методических подхода, которые хотя и не противоречат друг другу, но обычно являются альтернативными при разработке стандартов. В первом случае стандарт разрабатывается на объект стандартизации как на некую законченную систему (сущность). В другом стандарт создается на характеристики объекта. В чистом виде эти подходы не проявляются, но как правило, доминирует первый или второй подход при создании реального стандарта.

Оба подхода приемлемы, но для разных ситуаций. В том случае, когда объект стандартизации может быть представлен в виде статической модели (например, механическое устройство) предпочтительней первый подход с описанием стандарта как некой сущности.

При описании объекта стандартизации с помощью динамической модели (например, компьютерная технология, меняющаяся в зависимости от вычислительных ресурсов или от методов обработки) целесообразен параметрический подход, т.е. разработка стандарта с акцентом на качество результатов работы такого объекта.

В условиях быстрого обновления технологий стандарты первого типа не только быстро устаревают, но самое главное становятся препятствием для применения новых методов и технологий.

Стандарты второго типа допускают относительную модификацию объекта, главным в них является требование сохранения характеристик в процессе эксплуатации объекта. Другими словами, жизненный цикл у стандартов второго типа больше.

Область новых информационных технологий и геоинформатики характерна динамикой развития, обновлением технических средств, разработкой и применением новых подходов в программировании и обработке информации. По этой причине для геоинформатики целесообразней разработка стандартов на основе второго подхода.

#### **25.4. Стандартизация информационных продуктов и систем в международных организациях**

Для анализа стандартизации в области инструментальных программных средств (ИПС) целесообразно рассмотреть опыт международных организаций.

Ведущая международная организация по стандартизации ISO (the International Standards Organization). ISO является самой большой де-юре организацией стандартизации в области техники. В настоящее время, ISO состоит из 214 Технических Комитетов (TC) которые охватывают практически все технические направления: механику, электричество, химическую промышленность, авионавтику, нефтепродукты и т.д. [119]

Стандартизация в области информационных технологий обеспечивается объединенным техническим комитетом " JTC 1 ", в котором 26 стран участвуют непосредственно и 33 страны присутствуют в качестве наблюдателей.

JTC 1 был создан в 1987 как результат совместных действий ISO и IEC (the International Electrotechnical Commission). Основные области стандартизации НИТ обеспечиваются специализированными подкомитетами (SC) перечисленными в Таблице 25. 1

Таблица 25.1. Список подкомитетов объединенного технического комитета по стандартизации в области новых информационных технологий

JTC 1 / SGFS	Специальная группа функциональной стандартизации
JTC 1 / SC 1	Словари
JTC 1 / SC 2	Кодирование символов
JTC 1 / SC 6	Телекоммуникационный и информационный обмен между системами
JTC 1 / SC 7	Разработка программного обеспечения
JTC 1 / SC 11	Гибкие магнитные средства для цифрового обмена данных
JTC 1 / SC 7	Принципы организации элементарных данных
JTC 1 / SC 15	Объемы и структуры файлов
JTC 1 / SC 17	Идентификационные карты и связанные с ними устройства
JTC 1 / SC 18	Автоматизированная обработка и передача документов
JTC 1 / SC 21	Открытые системы связи, управление данными и открытая распределенная обработка
JTC 1 / SC 22	Языки программирования, их среды и интерфейсы программного обеспечения системы
JTC 1 / SC 23	Оптический дисковые картриджи для информационного обмена
JTC 1 / SC 24	Компьютерная графика и обработка изображения
JTC 1 / SC 25	Оборудование информационных технологий
JTC 1 / SC 26	Микропроцессорные системы
JTC 1 / SC 27	Методы защиты информационных технологий
JTC 1 / SC 28	Оборудование офиса
JTC 1 / SC 29	Кодирование аудио-, видео-, мультимедиа- и гипермедиа информации
JTC 1 / SC 30	Технологии открытого обмена электронными данными
JTC 1 / SC 31	Автоматизированный сбор данных



Другая ведущая организация ITU (International Telecommunications Union) - Международный телекоммуникационный союз, первоначально был создан национальными правительствами (в настоящее время их число свыше 184) для координации и разработки систем передачи данных во всем мире. В настоящее время ITU - это агентство Организации Объединенных Наций, основной функцией которого является обсуждение и разработка стандартов. Принципиальные цели ITU состоят в том, чтобы стимулировать продвижение технологии и гарантировать совместимость национальных сетей для эффективной организации международной связи.

В 1992 ITU был реорганизован: два его сегмента, CCITT и COIR были расформированы. Они были заменены тремя новыми отделами: телекоммуникационный сектор стандартизации (ITU-T), радиокommunikационный сектор (ITU-R), и телекоммуникационный сектор разработки (ITU-D). Негосударственные организации могут также быть членами сектора ITU с номером 380.

Список рекомендаций ITU-T для разных областей стандартизации ориентирован на связь. Применительно к НИТ следует выделить следующие серии: В-Средства выражения, G-Системы и средства передачи, К-Защита против несанкционированного доступа, Р-Качество передачи информации, Z-Языки программирования.

ETSI (the European Telecommunications Standards Institute) Европейский институт телекоммуникационных стандартов был создан в 1988 для определения европейских стандартов передачи данных и для сотрудничества с Европейским союзом радиовещания (EBU European Broadcasting Union) и CEN/CENELEC соответственно, в связанных полях радиовещания и технологиях электронного офиса.

В связи с изменением рынка передачи данных ETSI становится все большим экспертом при анализе рынка и прогнозом новых достижений. ETSI объединяет деятельность в области стандартизации с анализом поставщиков услуг, изготовителей, администрации, пользователей и исследованием взаимобмена.

Список рекомендаций технических комитетов ETSI областей стандартизации ориентирован на связь. Применительно к НИТ следует выделить следующие направления деятельности технического комитета TC11-Методы тестирования и спецификации.

Европейские организации стандартов: Comite Europeen Normalisation (CEN), Comite Europeen Normalisation Electrotechnique (CENELEC), являются ассоциациями, включающими национальные органы стандартизации и электротехнические комитеты следующих

Европейских стран: Австрия, Бельгия, Дания, Финляндия, Франция, Германия, Греция, Исландия, Ирландия, Италия, Люксембург, Нидерланды, Норвегия, Португалия, Швеция, Швейцария, Испания, и Великобритания.

Основная цель CEN состоит в планировании, проектировании и принятии Европейских стандартов как технической основы для общего гомогенного Европейского рынка. Поле CEN действия покрывает целый технический сектор за исключением электротехнической (CENELEC) и передачи данных (ETSI).

Для исследования в области НИТ в 1995 Европейскими Организациями Стандартизации, CEN, CENELEC и ETSI и др. был сформирован Отдел стандартизации информационных и коммуникационных технологий. Целью организации отдела явилось требование создания последовательной обоснованной программы стандартизации с учетом стремительного развития новых информационных технологий.

### **25.5. Особенности стандартизации инструментальных программных средств и программных продуктов**

Стандартизация в области информатизации имеет ряд существенных особенностей, отличающих ее от стандартизации в других областях, например в области машиностроения. Как показывает опыт, этого не понимает часть специалистов, работающих или начинающих работать в области автоматизации информационных процессов.

Стандартизация информационных программных средств и программных продуктов включает помимо традиционных методов оформления и подготовки стандартов специальные вопросы, имеющие отношение к новым информационным технологиям [96, 97, 119].

Обмен данными во многих технологиях осуществляется с применением моделей данных. По этой причине модели данных служат *первичной основой стандартизации* при создании программных средств и информационных продуктов, а следовательно, должны быть объектом изучения. Специально этим вопросом занимается объединенный технический подкомитет стандартизации JTC 1 / SC 14 ISO .

Стандарты качества продукции и качества информации входят в круг вопросов стандартизации информационных продуктов.

В тесной взаимосвязи с качеством информационных средств и продуктов находится надежность. Стандартизация информационных программных средств (ИПС) требует обеспечения надежности. Для

этого необходимо знать и использовать разнообразные методы оценки надежности программной продукции, чему посвящена деятельность объединенного технического подкомитета стандартизации JTC 1 / SC 7 ISO, серия P-ITU-T.

Одним из принципов стандартизации всех без исключения видов продукции является *системность*. Она определяет стандарт не как нечто отвлеченное, а как элемент системы и приводит к созданию систем стандартов, связанных между собой внутренней сущностью и концепцией объектов стандартизации.

Системность требует применения концептуального подхода при стандартизации и разработке современных информационных средств. Это достигается с применением методов системного анализа основанного на общей теории систем (ОТС) (подкомитет JTC 1 / SC 7 ISO).

Стандартизация разработки информационных средств, в отличие от стандартизации многих других видов продукции, требует лингвистического обеспечения и формального описания неких правил (синтаксиса), которым данное информационное средство должно подчиняться. Такая формализация может реализоваться с использованием разных формальных методов. Одним из примеров такого метода может служить формализм БНФ, который явился основой разработки самого первого стандарта алгоритмического языка Алгол.

Другим примером стандартизованного построения любой информационной модели является применение семиотического подхода при построении. Согласно этому подходу любая информационная модель должна включать три составные части: семантику, синтаксис, прагматику. Отсутствие любой из названных составляющих модели является признаком ее нестандартного построения.

Построение программ и их структур осуществляется разными путями. При этом возникает проблема сопоставимости и возможности обмена данными между современными программами и программами, разработанными ранее по другим стандартам. Поэтому стандартизация разработки современных информационных средств требует освещения всех основных подходов проектирования информационных и программных средств: модульного, объектно-ориентированного (объектного), структурного и др.

Особенностью разработки программных средств и информационных продуктов является их ориентация на группы конкретных пользователей. Это выражается в разработке специализированных интерфейсов. Поэтому стандартизация разработки информационных

средств, особенно на современном этапе, в отличие от стандартизации многих других видов продукции, требует создания удобного интерфейса пользователя. Главная задача интерфейса пользователя - обеспечение широкого сервиса при максимальной прозрачности информационного средства.

Стандартизация требует тестирования и контроля информационных средств [108]. В частности, при тестировании должны быть оценены основные модели типов воздействий на информационные и программные средства (ИПС) со стороны программного обеспечения (П), техники (Т), информации (И) и человека (Ч). В соответствии с этим при тестировании ИПС должны быть оценены соответствующие отказы и рассмотрены такие аспекты как: программный; аппаратный; информационный; эргатический.

## **25.6. Основные положения стандартизации**

Для эффективного обмена информацией между информационными системами и обеспечения качества информационных продуктов необходимы стандарты на структуру, содержание и протоколы передаваемой информации. В соответствии со стандартами ISO серии 9000 все элементы, требования и положения, принятые предприятием для общего руководства и систем качества, должны быть документированы.

Условиями, на которых основана система стандартов являются:

- наличие единой терминологической базы;
- системы классификации;
- системы кодирования объектов.

Стандарты основываются на обобщенных результатах науки, техники и практического опыта и направлены на обеспечение, управление и улучшение качества информационных процессов систем и продуктов. Они позволяют распределить ответственность, права и обязанности, установить порядок взаимодействия подразделений и исполнителей при выполнении своих функций.

Стандарт является нормативным документом, который определяет:

объект стандартизации  
цели и задачи стандартизации,  
основные принципы и организацию работ по стандартизации в Российской Федерации,  
категории (термины и определения) нормативных документов по стандартизации,  
виды стандартов,

основные положения по международному сотрудничеству Российской Федерации в области стандартизации, по применению стандартов и технических условий, по государственному надзору за соблюдением стандартов.

### **25.7. Объект стандартизации**

**Объектом стандартизации** является предмет (продукция, процесс, услуга), подлежащий или подвергшийся стандартизации. В области информационных продуктов под объектом стандартизации понимаются данные, алгоритмы и программы, результаты обработки (продукция), информационные процессы и услуги, методы проектирования и построения информационных процессов и систем, методы и технологии сбора, накопления, анализа, обработки и выдачи информации. Стандартизация может ограничиться определенными аспектами (свойствами) любого объекта. Например, применительно к информационным продуктам их надежность, качество, точность и полнота содержания могут быть стандартизированы отдельно.

### **25.8. Основные цели и задачи стандартизации**

Основными целями стандартизации являются:

- повышение качества продукции в соответствии с потребностями производства и потребления;

- защита интересов потребителей и государства в сфере обеспечения безопасности жизни, здоровья и имущества людей, охраны окружающей среды;

- обеспечение совместимости, взаимозаменяемости и модернизации продукции;

- экономия ресурсов, улучшение экономических показателей производства;

- обеспечение конкурентоспособности продукции на мировом рынке

Основными задачами стандартизации являются:

- обеспечение эффективности взаимодействия между разработчиками, заказчиками и потребителями;

- определение требований по совместимости и взаимозаменяемости информационной продукции;

- согласование характеристик продукции и ее элементов;

- унификация исходных данных и выходных продуктов;

- определение метрологических норм, правил;

- нормативно-техническое обеспечение контроля (испытаний, анализа, измерений) и оценки качества продукции;

- нормативно-техническое обеспечение сертификации продук-

ции;

унификация требований к технологическим процессам;  
создание и ведение систем классификации и кодирования информации;

нормативное обеспечение межгосударственных и государственных социально-экономических и научно-технических программ (проектов) и инфраструктурных комплексов;

создание системы информационного обеспечения выпускаемой продукции.

## 25.9. Категории стандартизации

Опорной категорией, применяемой при построении других, служит понятие стандартизации.

**Стандартизация** - это деятельность, направленная на достижение оптимальной степени упорядочения в определенной области посредством установления положений для всеобщего и многократного применения в отношении реально существующих или потенциальных задач.

Стандартизация *применяется* при разработке, опубликовании и применении стандартов и технических условий на продукцию. Важнейшими *результатами деятельности* по стандартизации являются повышение степени соответствия продукции, процессов и услуг их функциональному назначению, устранение барьеров в торговле и содействие научно-техническому и экономическому сотрудничеству.

**Стандарт** - нормативный документ по стандартизации, разработанный специалистами в данной области, согласованный по существенным вопросам у большинства заинтересованных сторон и утвержденный признанным органом (или предприятием). В стандарте устанавливаются правила, общие принципы и характеристики для всеобщего и многократного использования, объектов стандартизации.

**Технические условия** - нормативный документ на конкретную продукцию (услугу), утвержденный разработчиком по согласованию с заказчиком (потребителем).

## 25.10. Виды стандартов

**Стандарт предприятия** - (СТП) стандарт, утвержденный предприятием и применяемый только на данном предприятии.

**Отраслевой стандарт** - (ОСТ) стандарт, утвержденный министерством (ведомством) Российской Федерации.

**Национальный стандарт** - стандарт, принятый национальным

органом по стандартизации одной страны.

**Государственный стандарт Российской Федерации** - стандарт, утвержденный Государственным комитетом Российской Федерации по стандартизации, метрологии и сертификации (ГОСТ Р - Госстандарт России).

**Международный стандарт** - стандарт, принятый международной организацией по стандартизации.

**Региональный стандарт** - стандарт, принятый региональной международной организацией по стандартизации. К стандартам регионального типа относятся межгосударственные стандарты.

**Межгосударственный стандарт** - стандарт, принятый государством, присоединившимся к соглашению о проведении согласованной политики в области стандартизации, метрологии и сертификации и применяемый ими непосредственно.

**Комплекс стандартов (профиль)** - совокупность взаимосвязанных стандартов, объединенных общей целевой направленностью и устанавливающих согласованные требования к взаимосвязанным объектам стандартизации.

### 25.11. Жизненный цикл продукции

Как уже отмечалось, жизненный цикл является важнейшей характеристикой любой продукции, особенно в условиях рыночных отношений. Реальный жизненный цикл определяется спросом на продукцию или технологию и характеризует их конкурентоспособность.

**Жизненным циклом продукции** называют временной период подготовки, выпуска и продажи продукции (товара), который разбивают на несколько этапов, на каждом из которых продукцию надо по-разному рекламировать, продавать и менять цену.

Существует упрощенное и полное описание периодов жизненного цикла. При *упрощенном* описании выделяют четыре этапа жизненного цикла:

- внедрение,
- рост,
- зрелость,
- спад.

В методологии контроля качества, применяемой в нашей стране, ранее рассматривалось четыре стадии жизненного цикла продукции: исследование и разработка; изготовление; обращение и реализация; эксплуатация и потребление.

При полном описании жизненного цикла используют между-

народный стандарт ISO 9000-9004. В соответствии с этим стандартом жизненный цикл продукции (в ISO 9000-9004 он называется "петля качества") разделен на большее число этапов (рис.25.1):

- маркетинг, поиски и изучение рынка;
- проектирование и (или) разработка технических требований, разработка продукции;
- материально-техническое снабжение;
- подготовка и разработка производственных процессов;
- производство;
- контроль, проведение испытаний и обследований;
- упаковка и хранение;
- реализация и распределение продукции;
- монтаж и эксплуатация;
- техническая помощь и обслуживание;
- утилизация после использования.

### 25.12. Система качества

**Система качества** предназначена для обеспечения качества, контроля за его соблюдением и повышением качества продукции. Жизненный цикл продукции составляет основу системы качества. На рис. 25.1. показана модель системы качества. Согласно данной модели система качества должна охватывать все стадии жизненного цикла продукции.

Это определяет требования к основным этапам жизненного цикла продукции (петли качества), связанные с качеством продукции. Маркетинг, проектирование, тестирование и контроль - это наиболее ответственные этапы в петле качества. Кроме того, большое значение имеет *документационное обеспечение* системы качества.

Контроль системы качества включает два направления. Контроль организации работ и контроль самой продукции. Контроль может быть внутренним и внешним.

Контроль организации работ включать проверку:

- организационной структуры;
- административных и рабочих процедур;
- людских и материальных ресурсов, оборудования;
- рабочих участков, операций и процессов;
- документации, отчетов, ведения данных.





Рис. 25.1. Модель системы качества

Контроль самой продукции включает проверку:

- производимой продукции (с целью определения степени соответствия стандартам и техническим требованиям);
- технологических операций и процессов с применением продукции;
- соответствие достигаемых практических результатов при использовании продукции и ее рекламируемых характеристик.

Результаты проведенной проверок, заключения и рекомендации предоставляются в форме документов для рассмотрения на соответствующем уровне. На основе результатов проверки осуществляется анализ и оценка системы качества, которая включает:

- 1) результаты проверки различных элементов системы качества;
- 2) оценку эффективности системы качества по достижению целей, поставленных в области качества;
- 3) предложения по поддержанию в рабочем состоянии системы качества в соответствии с изменениями, вызванными новыми технологиями, концепциями качества, рыночной стратегией, социальными или окружающими условиями и производственной среды.

Анализ и оценка, включающие полученные результаты, выводы и рекомендации, служат основой для принятия решений по выпуску и модернизации продукции.

### 25.13. Качество информации

Информация как продукция и товар обладает качеством. Как ресурс и интеллектуальный потенциал она имеет особенности при оценке ее качества. Основными характеристиками качества информации, наряду с общепринятыми для других видов продукции, являются: репрезентативность, содержательность, прагматизм, достаточность, точность, актуальность, устойчивость.

**Репрезентативность** информации связана с правильностью ее отбора и формирования в целях адекватного отражения реальности или свойств объектов. Эта характеристика важна при выборочных наблюдениях или методах сбора данных.

**Содержательность** информации определяется либо коэффициентом информативности, т.е. отношением количества синтаксической информации к ее общему объему, либо коэффициентом содержательности отношением семантической информации к ее общему объему [29].

**Достаточность** (полнота) информации характеризует необхо-

димый набор данных для решения задач или принятия решений.

**Актуальность** информации характеризует ее временной аспект. Она определяется степенью ценности информации на момент использования. Эта характеристика качества особенно тем, что с течением времени актуальность информации снижается независимо от того, использовалась эта информация или нет.

**Точность** определяется степенью соответствия данных или параметров модели реальному состоянию процесса или объекта. Она оценивается чаще всего с помощью числовых мер.

**Прагматизм** определяется ценностью информации. Для его оценки используют целевую функцию. Поэтому ценность информации определяют в тех же самых единицах, в которых измеряется целевая функция. (денежная форма, производительность обработки, скорость передачи данных, время принятия решений, число операций за единицу времени и т.д.).

**Устойчивость** информации определяется ее способностью реагировать на случайные или санкционированные изменения данных без нарушения процессов обработки или информационного обмена. Более широко известно понятие помехоустойчивости, связанное с передачей кодированной информации. Устойчивость обеспечивается избыточностью данных, применением помехоустойчивых методов кодирования данных, системами защиты информации.

## **26. Сертификация информационных средств и программных продуктов**

Сертификация информационных средств тесно связана со стандартизацией и является гарантией соответствия информационной продукции существующим стандартам на момент ее выпуска.

Организация сертификации информационных продуктов и услуг возлагается на Госстандарт России, Роскоминформ, отраслевые руководящие органы в пределах компетенции, определенной для них в нормативных документах. Сертификация информационных средств и программных продуктов (ИПС) проводится одновременно с проверкой нормативной документации на данную продукцию стандартам или другим нормативным требованиям, включенным в перечень нормативной документации системы сертификации.

Кратко перечислим, что подлежит сертификации.

1) Сертификации подлежат защищенные технические, программно-технические, программные средства, системы, сети вычислительной техники и связи, средства защиты и средства контроля эффективности защиты;

2) Обязательной сертификации подлежат средства, в том числе и иностранного производства, предназначенные для обработки информации с ограниченным доступом и прежде всего содержащей сведения составляющие государственную тайну, а также использующиеся в управлении экологически опасными объектами, вооружением и военной техникой и средства их защиты. В остальных случаях сертификация носит добровольный характер и может осуществляться по инициативе производителя и потребителя;

Сертификация осуществляется в виде следующей последовательности этапов:

Разработка на основе ТЗ нормативного документа на продукцию в ранге Отраслевого стандарта (ОСТ) или Временных технических условий (ВТУ).

Разработка методики испытаний под утвержденный ОСТ или одновременно с ним.

Разработка инструментальных средств (тестов и стендов) в соответствии с ОСТ и методикой испытаний.

Включение ОСТ в область аккредитации.

Проведение испытаний на соответствие продукции требуемым нормам и стандартам.

Сертификат может быть выдан непосредственно изготовителем продукции (самосертификация) или сторонней организацией, которая проводит работы от испытаний до выдачи сертификата. Эта сторонняя организация оценивает и подтверждает соответствие продукции существующим стандартам.

## 26.1. Цели сертификации ИПС

Целями сертификации являются:

- содействие пользователям в компетентном выборе программного обеспечения;
- функциональная стандартизация программного обеспечения;
- улучшение качества программного обеспечения;
- защита пользователей от недобросовестности производителей программного обеспечения;
- подтверждение показателей качества программного обеспечения, заявленных его изготовителями.

Сертификация проводится в соответствии с Федеральным законом "О защите прав потребителя", Федеральным законом "О сертификации продукции и услуг", Постановлением Госстандарта РФ "Об утверждении "Порядка проведения сертификации в Российской Фе-

дерации". Любая организация, разработавшая или эксплуатирующая программное обеспечение, согласная с Процедурой сертификации программного обеспечения (далее - Процедурой) и Требованиями, на соответствие которым проводится сертификация, вправе подать заявку на сертификацию.

## 26.2. Заявка на сертификацию

Сертификация может быть осуществлена по заявленному требованию. В этом случае заявитель просит третью сторону подтвердить, что его продукция обладает теми свойствами, которые он регламентирует. В этом случае он может быть *поставщиком* продукции.

Заявка на сертификацию подается от имени организации, разработавшей или эксплуатирующей систему управления базами данных или информационно-поисковую систему (далее - организация-заявитель).

Заявитель может быть *покупателем* продукции. Он может требовать подтверждения того, что приобретаемый продукт соответствует заявляемым спецификациям, а у поставщика нет сертификата соответствия.

Данный случай ориентирован на конкретный вид продукции. В этом случае орган сертификации устанавливает требование к конкретному виду продукции и выдает сертификат соответствия.

Во втором случае сертификация включает разработку ОСТа на продукцию, составлению ТУ на испытания, разработку инструментальных средств, проведение испытаний по правилам сертификации.

Заявка на сертификацию включает основную информацию об организации-заявителе, разработчике программного обеспечения, программном обеспечении. В случае указания в заявке неточных, неполных или недостоверных данных орган сертификации имеет право отклонить заявку.

## 26.3. Условия сертификации

На сертификацию принимаются только рабочие версии ИПС. Тестирование проводится в среде, для которой информационный продукт создается. Для участия в сертификации необходимо предоставить в орган сертификации примерно следующие документы и материалы:

- заявку на сертификацию программного обеспечения;
- письменное подтверждение согласия с Процедурой и Требованиями сертификации, подписанное руководителем организации-заявителя;

- письменное согласие от организации, разработавшей программное обеспечение (в случае, если заявка подана от имени организации эксплуатирующей программное обеспечение);
- письменное подтверждение согласия на публикацию результатов тестирования и основных характеристик программ, в случае положительного результата сертификации;
- две дистрибутивные копии программного обеспечения на дискетах 3,5 дюйма (в формате, согласованном со специалистами органа сертификации). Дискеты должны сопровождаться отпечатанным списком файлов, записанных на них, в списке должны быть указаны объем файлов, дата и время их создания;
- копию документа (лицензию), подтверждающего легальность покупки фирмой-разработчиком программных средств, использованных для разработки программного обеспечения, предоставленного на сертификацию;
- один комплект документации на программное обеспечение (в бумажной форме);
- в случае если заявка подана повторно, она должна сопровождаться списком исправленных ошибок со ссылкой на номера отчетов об обнаруженных ошибках;
- в случае, если ОС распространяется заявителем среди организаций, осуществляющих экономическую деятельность, то организация-заявитель должна предоставить список пользователей, зарегистрированных на дату подачи заявки (для каждого пользователя ОС необходимо указать: полное официальное наименование, место нахождения, почтовый адрес, телефон, факс). Это выполняется для того, чтобы в случае обнаружения дефектов в сертифицируемой продукции и последующем их устранении заявителем, орган сертификации не нес ответственности за эксплуатацию дефектных ИПС.

Документы и материалы, предоставленные организацией-заявителем на сертификацию, не возвращаются.

В случае неудачной попытки пройти сертификацию, организация-заявитель имеет право предоставить повторную заявку на сертификацию. Повторная заявка на сертификацию подается не ранее чем через 1 месяц после получения заявителем уведомления о несоответствии ИПС предъявляемым требованиям.

Улучшение качества ОС, в рамках сертифицированной версии, может не потребовать повторной сертификации. В этом случае орга-

низация-заявитель должна предоставить в орган сертификации перечень изменений, внесенных в сертифицированную версию, который рассматривается на заседании орган сертификации.

В зависимости от характера и объема внесенных изменений, орган сертификации принимает решение о необходимости повторной сертификации ОС. Решение направляется организации-заявителю в 2-х недельный срок с момента подачи перечня изменений, внесенных в сертифицированную версию.

#### **26.4. Процесс сертификации**

Тестирование объекта сертификации выполняется специалистами испытательной лаборатории. По окончании процесса тестирования его результаты (списки отчетов об обнаруженных ошибках, отчеты об обнаруженных ошибках), передаются в орган сертификации, который принимает решение о выдаче сертификата программному обеспечению, либо об отказе в выдаче такового.

После начала процесса сертификации исправления к ИПС и/или документации не принимаются.

#### **26.5. Результаты сертификации**

В случае положительного результата сертификации, организации-заявителю выдается сертификат, подтверждающий, что предоставленный на сертификацию информационный продукт прошел тестирование и соответствует Требованиям. В этом случае результаты сертификации и основные характеристики ОС могут публиковаться в средствах массовой информации.

Действие сертификата распространяется только на версию (с указанием номера и даты выпуска версии), предоставленную на сертификацию.

Орган сертификации несет ответственность за соответствие сертифицированной версии программного обеспечения (предоставленной на сертификацию) Требованиям. В случае наличия в программном обеспечении ОС дополнительных функций и операций, не предусмотренных Требованиями, органа сертификации вправе не осуществлять тестирование этих функций и операций и не несет ответственности за их корректность.

Программное обеспечение ОС, прошедшее сертификацию, **не маркируется** знаком соответствия.

В случае отрицательного результата сертификации, организация-заявитель уведомляется об этом органом сертификации. Организации-заявителю предоставляются результаты сертификации (выписка из протокола соответствующего решения орган сертификации,

списки отчетов об обнаруженных ошибках, отчеты об обнаруженных ошибках).

При несогласии организации-заявителя с результатами сертификации, он имеет право подать апелляцию в орган сертификации.

Апелляции рассматриваются на заседании орган сертификации по мере их поступления с участием представителей организации-заявителя и сотрудника органа сертификации, проводившего тестирование программного обеспечения.

Решение по итогам рассмотрения апелляции принимается руководством органа сертификации. Организация-заявитель уведомляется об итогах рассмотрения апелляции органом сертификации.

### **26.6. Оплата работ по сертификации**

Оплата работ по сертификации производится заявителем в порядке, установленном в соответствии с заключенным соглашением между органом сертификации и заявителем и прилагаемым расчетом стоимости работ.

Стоимость работ по сертификации определяется на основе расчета в зависимости от вида и объема работ по сертификации и устанавливается соглашением сторон. При этом уровень рентабельности работ не должен превышать 50%.

Сумма средств, израсходованных заявителем на проведение сертификации, относится на себестоимость сертифицируемых средств. Оплата осуществляется до начала выполнения работ по сертификации и не зависит от результатов сертификации.

### **26.7. Срок действия сертификата**

Прекращение действия сертификатов определяется решением руководства органа сертификации на основе результатов проверки инспекционного контроля или при окончании установленного срока действия сертификата. Решение органа сертификации о прекращении действия ранее выданных сертификатов доводится до заинтересованных лиц в письменной форме в срок, не позднее семи рабочих дней с момента его принятия.



## Список сокращений

- АИИС - Автоматизированная интегрированная информационная система
- АИС - Аналитическая информационная система
- АС - Автоматизированная система
- АСИС - Автоматизированная справочно-информационная система
- АСНИ - Автоматизированная система научных исследований
- АСУ - Автоматизированная система управления
- БД - База данных
- ГИС - Геоинформационная система
- ГИС - Географическая информационная система
- ГОСТ Р - Государственный стандарт Российской Федерации
- ГСС - Государственная система стандартизации Российской Федерации
- ЗИС - Земельная информационная система
- ДДЗ - Данные дистанционного зондирования
- ИМ - Иерархическая модель
- ИСЗ - Искусственный спутник Земли
- ИПС - Информационное программное средство
- КН - Кадастровый номер
- КТС - комплекс технических средств
- ЛВС - Локальная вычислительная сеть.
- ЛПР - Лицо принимающее решение
- НСД - Несанкционированный доступ ( к информации)
- ОСТ - Отраслевой стандарт
- ОС - Операционная система
- ОУ - Объект управления
- ПО - Программное обеспечение
- РМ - Реляционная модель
- САПР - Система автоматизированного проектирования
- СВТ - Средства вычислительной техники
- СЗИ НДС – Система защиты информации от несанкционированного доступа.
- СУ - Система управления
- СУБД - Система управления базами данных
- СТП - Стандарт предприятия
- ТЗ - Техническое задание

ТСЗ - Технические средства защиты

ТСО - Технические средства обнаружения

ЦМК - Цифровая модель карты

ЦММ - Цифровая модель местности

ЦМО - Цифровая модель объекта

ЦМР - Цифровая модель рельефа

ЦМЯ - Цифровая модель явления

ЭК - Электронная карта

ЭС - Экспертная система

CASE - (Computer Aided System Engineering) - Технология "автоматизированного построения (информационной) системы"

CEN (Comite Europeen Normalisation) - Европейский комитет по нормализации стандартов

DCW - Digital Chart of the World - цифровая карта мира (суши) в формате ArcInfo

ER - Модель "Сущность-связь"

ESRI - Институт исследования систем окружающей Среды (США)

ETSI (the European Telecommunications Standards Institute) Европейский институт телекоммуникационных стандартов

GPS - (Global Position System) - Глобальная система позиционирования (местопределения)

HTML (HyperText Markup Language) - язык разметки гипертекста, используется при подготовке документов для системы WWW

ISO (the International Standards Organization) 1. Международная организация по стандартизации. 2. Серия международных стандартов

ITU (International Telecommunications Union) - Международный телекоммуникационный союз вторая по значимости международная организация по стандартизации

OGIS - Open Geodata Interoperability Specification - Открытая спецификация обмена геоданными,

TCP/IP (Transport Control Protocol / Internet Protocol) - протокол передачи данных по сети Интернет

TIN - Треугольные сети неправильной формы

TSR (Terminate and Stay Resident) - Резидентная программа

UTM - Универсальная поперечно-цилиндрическая проекция Меркатора (проекция Гаусса-Боагта)

URL (Uniform Resource Locator) - универсальный указатель ресурсов).

## Терминологический словарь

**Абстракция** - процедура структуризации (типизации) данных. Различают два вида абстракции: *обобщение* и *агрегация*.

**Автоматизированная справочно-информационная система (АСИС)** - автоматизированная система обслуживания потребителей информацией в удобной для восприятия и переработке форме.

**АСИС документографическая** - автоматизированная информационная система, информационным массивом которой служат различные неформализованные (слабо типизированные данные) документы (цитаты, статьи, письма и т. д.) на естественном или ограниченном искусственном языке.

**АСИС фактографическая** - автоматизированная информационная система, информационный массив которой составляется из формализованных записей (сильно типизированных данных).

**Агрегативные данные** - набор данных для формирования объекта из его частей на основе процедур *агрегации*.

**Агрегация** - процедура структуризации данных. Заключается в конструировании объекта из других базовых объектов на основе чего создается агрегативная модель. Соотносится с понятием "*есть некоторое...*".

**Атрибут** - элементарное данное, описывающее свойства *сущностей*.

**Атрибутивные данные** - данные с присвоенной им описательной информацией (время, тема). Являются вторым основным классом данных в ГИС, которые дополняют *координатные данные* необходимыми описаниями.

**Большой круг** - воображаемый круг на земной поверхности, образованный плоскостью, проходящей через центр земного шара.

**Буферные зоны** - зоны, границы которых удалены на известное расстояние от любого объекта на карте. Буферные зоны различной ширины могут быть созданы вокруг выбранных объектов на базе таблиц сопряженных характеристик.

**Валентность** - специальная (топологическая) характеристика данных *вершины (узла)* сети, определяемая количеством *звеньев* в нем. Концы обособленных линий "одновалентны". Для уличных сетей (пресечения типа "крест") наиболее характерны четырехвалентные узлы. В гидрографии чаще всего встречаются трехвалентные узлы.

**Векторизация** - процедура (ручная или автоматизированная) выделения векторных объектов с растрового изображения и получения их в векторном формате.

**Векторные (объектные) модели ГИС** - модели данных ГИС, строящиеся на векторах, занимающие часть пространства в отличие от занимающих все пространство *растровых моделей*. При построении векторных моделей создаются целостные объекты путем соединения точек прямыми линиями, дугами окружностей, полилиниями, поэтому векторные модели называют также **объектными**. Основные части векторной модели: геометрические объекты (точки, линии и полигоны); атрибуты (признаки, связанные с объектами); связи между объектами.

**Верификация** - процедура проверки. Может осуществляться автоматически или с помощью эвристических методов, вызываемых явной директивой пользователя или запускаемых при возникновении специфических ситуаций.

**Вершина (узел)** - соединение, начало или конец отрезка линии (*дуги, звена*).

**Внешний ключ** - столбец в реляционной модели, поддерживающий связь между разными реляционными таблицами.

**Внутренний идентификатор пространственного объекта** - целое число, являющееся служебным идентификатором системы (уникальное для каждого объекта данного покрытия и назначаемое автоматически в процессе работы редактора). Может изменяться системой в процессе работы.

**Внутренний ключ** - столбец в реляционной модели, однозначно определяющий каждую строку таблицы.

**Генерализация** - в ГИС представляет собой набор процедур классификации и обобщения, предназначенных для отбора и отображения картографических объектов соответственно масштабу, содержанию и тематической направленности создаваемой цифровой карты. Относительно *информационного моделирования* генерализация может быть рассмотрена как группа методов, позволяющих сохранить информативность даже при уменьшении объема данных.

**Геоинформационная система** - автоматизированная информационная система, предназначенная для обработки пространственно-временных данных, основой интеграции которых служит географическая информация.

**Геокодирование** - соединение табличных данных адресных файлов с географическим положением объектов.

**Геоид** - общая фигура Земли, образованная поверхностью, которая совпадает с поверхностью Мирового океана в состоянии покоя и равновесия и продолжена под материками.

**Граф** - графическая схема, включающая основные элементы графа, называемые *вершинами (узлами)* и *звеньями (ребрами, дугами)*. В отличие от произвольно нарисованной схемы *графовая форма моделей*, как и *табличная форма представления моделей*, строится по определенным правилам.

**Граф неориентированный** - граф, в котором ребра неориентированы, т.е. возможен путь от одной вершины к другой в обоих направлениях.

**Граф ориентированный** - граф, в котором каждое ребро ориентировано, т.е. определен путь от одной вершины к другой. Примером служит вектор в аналитической геометрии.

**Графовая форма представления моделей** - форма, которая дает представление модели или ее характеристик в виде одного или совокупности взаимосвязанных графов.

**Долгота** - угол в плоскости экватора между меридианом точки и главным (нулевым) меридианом, проходящим через Гринвич (Англия). Она изменяется от  $-180^\circ$  (западная долгота) до  $+180^\circ$  (восточная долгота).

**Запись данных** - формализованное представление сложной информационной модели без описания ее структуры. Запись может быть логической и физической.

**Запись логическая** - информационная единица, соответствующая одному шагу обработки информации.

**Запись физическая** - порция информации, которая является единицей обмена данными между внутренней и внешней памятью ЭВМ.

**Звено (ребро, дуга)** - линейный объект ГИС, отрезок, соединяющий две точки отрезка (вершины, узлы).

**Знак** - элементарная единица информации, являющаяся реализацией свойств объекта в заранее заданной, структурно организованной знаковой системе. Знак можно рассматривать как реализацию *типа*, тип - как обобщение совокупности знаков. Следовательно, знак представляет индивидуальные свойства модели, а тип - ее общие свойства.

**Зона (растровой модели)** - включает соседствующие друг с другом ячейки, имеющие одинаковое значение. Зоной могут

быть отдельные объекты, природные явления, ареалы типов почв, элементы гидрографии и т.п.

**Изоколы** - линии, соединяющие на карте точки с одинаковыми значениями искажений, обусловленных свойствами картографической проекции. Форма изоколов зависит от вида проекции. В конических проекциях изоколы имеют форму дуг, в азимутальных - концентрических окружностей, в цилиндрических - параллельных прямых.

**Иерархическая модель** - модель структуры процесса или системы, структурная схема которой является ориентированным графом. Состоит из совокупности исходных и порожденных типов записей (уровней), связанных одной дугой или несколькими дугами. Допускает связи уровней типа один ко многим (1: N) или один к одному (1:1).

**Иерархический путь** (маршрутизация) - последовательность типов записей иерархической модели, начинающаяся с корневой записи, в которой типы записей выступают попеременно в ролях исходного и порожденного.

**Кадастровые карты** - графические отображения юридических описаний земельных участков.

**Квадратомическое дерево** - иерархическая модель со связью 1:4. В этой структуре исходный уровень рекурсивно подразделяется на четыре порожденных уровня.

**Классификация** - процедура абстракции (структуризации) данных. Состоит в соотнесении множества знаков одному типу.

**Координатные данные** (место) - один из основных классов геоинформационных данных, используется для указания местоположения исследуемого объекта на земной поверхности.

**Малый круг** - воображаемый круг на земной поверхности, образованный плоскостью, не проходящей через центр земного шара.

**Меридиан** - линия пересечения плоскостей географических меридианов с земной поверхностью.

**Метаданные** - дополнительные данные о данных, например, описатели таблиц, столбцов и т.д. Обычно хранятся в словаре данных (data dictionary).

**Модели аналоговые** - модели, создаваемые на основе физического или математического моделирования (аналитического описания), например цифровая модель рельефа, построенная на основе аналитического описания поверхности.

**Модели динамические** - модели, служащие для описания изменения процессов и моделей. Допускают изменение параметров и структур во времени.

**Модели дискретные** - модели, основанные на замене непрерывных функций набором дискретных значений аргументов и функций. Дискретность определяется шагом квантования. Для необходимости сохранения информативности дискретной модели по отношению к объекту шаг квантования должен выбираться с учетом теоремы Шенона-Котельникова.

**Модели сильно типизированные** - модели, в которых большинство данных удовлетворяет неким условиям и ограничениям и может быть отнесено к узкому подклассу (*типу*).

**Модели слабо типизированные** - модели, в которых данные разнородны по формату, структуре и слабо связаны условиями относительно известных *типов*.

**Модели статические** - модели, инвариантные относительно времени.

**Моделирование инвариантное** - моделирование, основанное на работе с полностью или частично унифицированными информационными элементами или структурами. Имеется возможность использования групповых операций, что повышает производительность труда по сравнению с индивидуальным моделированием. Инвариантность предполагает использование общих свойств моделируемых объектов (свойств типов или классов) безотносительно к индивидуальным характеристикам отдельных объектов. Примером могут служить оверлейные процедуры.

**Моделирование информационное** - моделирование, связанное с созданием и преобразованием разных форм информации, например графической или текстовой, в вид, задаваемый пользователем. В современных информационных системах оно осуществляется путем создания подсистемы документационного обеспечения.

**Моделирование семантическое** - моделирование, связанное с задачами кодирования и лингвистического обеспечения. В ГИС доля семантического моделирования велика на стадии сбора информации, что обусловлено большим объемом и разнообразием входной информации, сложностью ее структуры, возможным наличием ошибок.

**Моделирование эвристическое** - моделирование, основанное на интерактивной обработке, применяется при необходимости экспертных решений, при учете дуальных свойств объектов на видео-

изображениях и при решении специальных нетиповых задач. В ГИС и САПР оно выражается в том, что общение пользователя с ЭВМ осуществляется на основе сценария, учитывающего, с одной стороны, технологические особенности программного обеспечения, с другой - особенности и опыт обработки данной категории объектов.

**Модель "сущность-связь"** (ER-модель (Entity Relationship Model)) - графовая модель, дающая представление о предметной области в виде объектов, называемых сущностями, между которыми фиксируются связи. Для каждой связи определено число связываемых ею объектов. На схеме сущности изображаются прямоугольниками, связи - ромбами. Число связываемых объектов указывается цифрой на линии соединения объекта и связи. В ER-модели допускаются представления связей типа "один к одному" (1:1), "один ко многим" (1 : M) , "многие к одному" (M : 1), "многие ко многим" (M : N).

**Номенклатура** - система обозначений отдельных листов топографических карт.

**Обобщение** - процедура *абстракции* (структуризации) данных, соотносится с понятием: "есть часть...". Обобщение акцентирует сходство объектов, абстрагируясь от различий. Подразделяется на две категории: *собственно обобщение* и *классификацию*.

**Оверлейные операции** - процедуры наложения разноименных словес ( двух или более, при многократном повторении операции попарного наложения) с генерацией производных объектов, возникающих при их геометрическом наложении и наследовании их атрибутов. Наиболее распространен случай оверлея двух полигональных слоев.

**Параллель** - линия, образованная пересечением плоскости, параллельной плоскости земного экватора, с поверхностью Земли.

**Первичный атрибут отношения** - *атрибут*, присутствующий по крайней мере в одном ключе, все другие атрибуты не первичные.

**Первичный ключ** - столбец (или комбинация столбцов ), значения в которых однозначно идентифицируют каждую строку *реляционной модели*.

**Пиксель (ячейка)** - минимальный элемент растровой модели, двухмерный элемент пространства.

**Пилот-проект** - усеченный вариант реализации системы ( не только ГИС, но и любой другой), требует минимального риска вложения денег от заказчика. Выполняется либо без затрат со стороны



администрации (заказчика), либо в ограниченном объеме финансирования (2 - 10 % от полной стоимости проекта). Реализует ограниченное количество функций и служит демонстрационным вариантом полной ГИС.

**Плоскость географического (астрономического) меридиана** - плоскость, проходящая через ось вращения Земли и отвесную линию в точке земной поверхности.

**Плоскость земного экватора** - плоскость, проходящая через центр Земли перпендикулярно к оси вращения.

**Позиционная точность координатных данных** - степень отклонения данных ГИС о местоположении объекта от его истинного положения на местности.

**Полигон** - пространственная область, ограниченная упорядоченным набором связанных дуг, которые образуют замкнутый контур.

**Пользовательский идентификатор (внутренний ключ)** пространственного объекта - целое число, предназначенное для связи объектов цифровой карты с базой (таблицами) тематических данных, назначаемое пользователем.

**Пошаговая детализация** - процедура, обратная *агрегации*. Применяется для разбиения агрегативной модели на составляющие ее части.

**Предметная область** - подмножество (часть реального мира), на котором определяется набор данных и методов манипулирования с ними для решения конкретных задач или исследований.

**Примитив** - тип *векторной модели*, атомарная графическая модель векторизации. Например, могут использоваться следующие типы векторных моделей: линия, дуга, текст, полилиния, контур и размерная линия.

**Проекции произвольные** - проекции, которые имеют искажения углов, площадей и длин, но эти искажения распределены по карте например так, что минимальные искажения имеются в центральной части и возрастают к краям. Среди произвольных проекций выделяют **равнопромежуточные** (*equidistant projection*), в которых искажения длин отсутствуют по одному из направлений: вдоль меридиана или вдоль параллели.

**Проекции равновеликие** - проекции, которые не искажают площадей, но искажают углы и формы объектов.

**Проекции равноугольные** - проекции, в которых сохраняются без искажений углы и формы малых объектов, но резко деформируются длины и площади объектов.

**Проекционные преобразования** - группа математических процедур ГИС, осуществляющих переход от одной картографической проекции к другой или от пространственной системы к картографической проекции.

**Псевдоузел** - узел, принадлежащий только двум дугам либо одной замкнутой дуге, у которой начальная и конечная вершины совпадают.

**Разграфка** - разделение топографических карт на листы .

**Растровая модель (растр) ГИС** - *дискретная* модель , в которой в качестве атомарной используют двухмерный элемент пространства - *пиксель (ячейка)*. Упорядоченная совокупность атомарных моделей образует растр, который в свою очередь, является моделью карты или геообъекта. В растровых моделях дискретизация осуществляется путем отображения геообъектов в пространственные ячейки, образующие регулярную сеть. При этом каждой ячейке соответствует одинаковый по размерам, но разный по характеристикам (цвет, плотность) участок поверхности объекта. В ячейке модели содержится одно значение, усредняющее характеристику участка поверхности объекта. В теории обработки изображений эта процедура известна под названием пикселизация.

Если *векторная модель* дает информацию о том, где расположен тот или иной объект, то растровая модель дает информацию о том, что расположено в той или иной точке территории. Это определяет основное назначение растровых моделей - непрерывное отображение поверхности. Основные характеристики растровых моделей:

**Значение** - элемент информации, хранящийся в элементе растра (пикселе).

**Ориентация** - угол между направлением на север и положением колонок растра.

**Разрешение** - минимальный линейный размер наименьшего участка пространства (поверхности), отображаемый одним пикселем. Более высоким разрешением обладает растр с меньшим размером ячеек. Высокое разрешение подразумевает обилие деталей, множество ячеек, мелкий размер ячейки.

**Реляционная модель** - табличная модель данных, основным средством структуризации в которой является отношение. Таблица состоит из строк и столбцов и имеет имя, уникальное внутри базы данных. Таблица отражает тип объекта реального мира (*сущность*), каждая ее строка - конкретный объект, каждый столбец - *атрибут*.

*Сущности, атрибуты* и связи хранятся в таблицах как данные определенного *типа*. В реляционных моделях данных не предусматривается поддержание логической упорядоченности, однако, кортежи помещаются в физическую память в соответствии с некоторым порядком. Физическая упорядоченность используется для выборки.

**Референц-эллипсоид** - эллипсоид, применяемый в качестве математической модели Земли.

**Слой** - покрытие, определяемое его тематической определенностью (растительность, рельеф, административное деление и т.д.) и координатной принадлежностью (точечный, линейный, полигональный).

**Собственно обобщение** - процедура соотнесения множества типов одному типу.

**Специализация** - процедура (обратная *обобщению*) порождения типов на основе общего класса типов.

**Сущность** - элемент модели (совокупность *атрибутов* и *знаков*), описывающий законченный объект или понятие.

**Табличная форма представления моделей** - форма, которая дает представление модели или ее характеристик в виде одной или совокупности взаимосвязанных таблиц. При этом данные в ячейках таблицы не могут заноситься произвольно, они подчиняются определенным правилам, в частности по столбцам располагают типизированные данные.

**Территориальные уровни использования ГИС:**

**Глобальный уровень** - Россия на глобальном и евразийском фоне масштаб, 1 : 45 000 000 - 1 : 100 000 000;

**Всероссийский уровень** - вся территория страны, включая прибрежные акватории и приграничные районы, масштаб 1 : 2 500 000 - 1 : 20 000 000;

**Региональный уровень** - крупные и природные экономические регионы, субъекты федерации, масштаб 1 : 500 000 - 1 : 4 000 000;

**Локальный уровень** - области, районы, национальные парки, ареал кризисных ситуаций, масштаб 1 : 50 000 - 1 000 000;

**Муниципальный уровень** - города, городские районы, пригородные зоны, масштаб 1 : 50 000 и крупнее.

**Тип** - Совокупность моделей или объектов, объединенная общим набором признаков или класс подобных *знаков*.

**Типизация** - процедура объединения данных по набору заданных признаков или выделения из множества данных тех, которые удовлетворяют заданным критериям (или признакам).

**Тип записи иерархической модели** - вершина в графовой форме иерархической модели, соответствующая типу сущности.

**Типы координатных данных в ГИС:** точки (узлы, вершины); линии (дуги, звенья); полигоны (районы, ареалы).

**Топологические свойства (фигур)** - свойства, которые не изменяются при любых деформациях, производимых без разрывов или соединений.

**Связанность векторов.** Контур, дороги и прочие векторы должны храниться как взаимосвязанные объекты, а не независимо;

**Связанность и примыкание районов.** Информация о взаимном расположении районов и об узлах пресечения районов;

**Пересечение.** Информация о типах пересечений (*валентности*) позволяет воспроизводить мосты и дорожные пересечения;

**Близость.** Показатель пространственной близости линейных или ареальных объектов.

**Топологическое описание (объектов ГИС)** - описание, которое определяется связанностью и взаимным расположением линейных объектов.

**Трансформирование** - разновидность процедур проекционных преобразований. Применяется для преобразования координат точек снимка в другую проекцию. Частным случаем является преобразование плоского изображения в плоское с целью устранения перспективных искажений.

**Уровенная поверхность** - поверхность, образующая тело, ограниченное поверхностью воды океанов. Основное ее свойство заключается в том, что потенциал силы тяжести на ней имеет одно и то же значение. Другими словами эта поверхность везде перпендикулярна отвесной линии, т.е. везде горизонтальна.

**Файл** - порция информации, имеющая начало и конец, хранящаяся на машинных носителях. Может содержать данные или программу.

**Физическая модель ГИС** - модель среды хранения данных физического уровня. Строится с учетом реальных инструментальных систем и на их основе. Может рассматриваться как реализация *даталогической модели*.

**Цифровая модель местности (ЦММ)** - комплексная модель местности, которая должна содержать четыре основных свойства, вытекающие из ее определения, а именно:

Как **цифровая** она должна быть оптимально организована и удобна при работе на ЭВМ. Это означает, что для полной ее реализации должна быть определена ее "физическая" структура.

Как **модель** вообще ЦММ должна быть определена на известном классе моделей. Это означает, что она должна иметь вполне определенную структуру и содержать в своей основе одну из базовых моделей данных, а также удовлетворять требованиям и обладать общими свойствами моделей соответствующего класса безотносительно к *предметной области*.

Как **модель местности** - она должна содержать специальную информацию данной конкретной предметной области. Это означает, что ЦММ должна содержать элементы координатного и атрибутивного описания, характеризующие как саму предметную область, так и индивидуальные свойства моделируемых объектов.

Как **структура базы данных** ЦММ должна иметь возможность для моделирования, многократного использования, анализа и решения различных задач. Для возможности многократного применения ЦММ должна быть переопределена (более информативна) по сравнению с информационной моделью ручной технологии, обеспечивающей получение разового продукта. Это требует выполнения обобщенного описания цифровых моделей местности на уровне *типов*, т.е. для этого необходимы предварительный анализ и последующая максимальная *типизация* пространственных объектов.

**Цифровое моделирование рельефа** - построение дискретной модели, основанное на переходе от аналоговой модели непрерывной поверхности (рельефа) к дискретной модели набора точек, оптимально отображающей форму этой поверхности.

**Цифровая фотокамера** - фотокамера, в которой изображение фиксируется не на фотопленку или фотопластинки, а на цифровой элемент (матричный или линейный). Это дает возможность хранить изображения в цифровом виде и записывать их непосредственно в компьютер. По точностным характеристикам соответствует либо любительским, либо метрическим камерам.

**Широта** - угол между точкой и экватором вдоль меридиана. Изменяется от  $-90^\circ$  (южный полюс) до  $+90^\circ$  (северный полюс).

**Экземплярция** - процедура (обратная *классификации*) порождения реализация на основе известной классификации.

**Экспертные системы (ЭС)** - класс автоматизированных информационных систем, содержащих базы данных и базы знаний, способных осуществлять анализ и коррекцию данных независимо от санкции пользователя, анализировать и принимать решения как по запросу, так и независимо от запроса пользователя и выполнять ряд аналитически - классификационных задач.

**ЭС диагностирующие** - системы, в качестве цели имеющие способность находить причины аномальности наблюдаемых явлений. Основой для анализа служат наборы данных, с помощью которых выявляются отклонения от эталонного поведения и в результате ставится диагноз.

**ЭС планирующие** - системы, предназначенные для выработки программы действий, необходимых для достижения определенных целей.

**ЭС прогнозирующие** - системы, предназначенные для построения сценария будущего. Основываясь на событиях прошлого и настоящего, они способны выводить вероятные следствия из заданных ситуаций. Для этой цели в прогнозирующих ЭС используются динамические параметрические модели.

## Литература

1. Артеменко В.В., Лойко П.Ф., Огарков А.П., Севостьянов А. В. "Кадастр земель населенных пунктов" – М.: Колос, 1997
2. Белов С.Б., Бобков В.А., Крюков К.А., Морозов М.А. Распознавание условных знаков // Информационные технологии. 2000, №7, с. 37-42.
3. Берж К. Теория графов и ее применение. - М.: Из-во иностр. лит., 1962
4. Берлянт А.М. Геоинформатика. - М.: МГУ, РАЕН, 1996-208с.
5. Блауберг И.В., Юдин Э.Г. Становление и сущность системного подхода. - М.: Наука, 1973.
6. Богатырь Б.Н., Гуриев М.А., Игнатъев М.Б., Лобанов В.С., Макарова Н.В., Осовецкий Л.Г., Советов Б.Я., Чижев С.А. Иванников А.Д. Концепция системной интеграции информационных технологий в высшей школе, Москва, РосНИИИС, 1993.- 72 с.
7. Бугаевский Л.М. Математическая картография. - М.: Златоуст, 1988. -480 с.
8. Бугаевский Л., М., Цветков В.Я. Геоинформационные системы. - М.: Златоуст, 2000. 224 с.
9. Бугаевский Л.М, Цветков В.Я., Флейс М.Э Терминологическая основа и вопросы обучения ГИС // Информационные технологии, 2000, №11, с.11-16
10. Бухтояров С.Г., Иванников А. Д., Ознобихин И.В. Устройство для моделирования цифровых объектов. А.С. СССР 1312601 от 22.01.87 г.
11. Виер С.Х. "Обработка картографических данных: земельные информационные системы. Системный подход к разработке земельных информационных систем./ в кн. Картография. Вып. 4. Геоинформационные системы – М: "Картоцентр-Геозидат",1994. С.112-124.
12. Генике А.А., Побединский Г.Г. Глобальная спутниковая система определения местоположения и ее применение в геодезии. М.: Москва, "Картоцентр-Геозидат", 1999
13. Гинзбург Г.А., Салманова Т.Д. Пособие по математической картографии. Тр. ЦНИИГАиК, Вып. 160 - М.: ЦНИИГАиК, 1964.
14. Гладкий В.А. Кадастровые работы в городах. Новосибирск, Наука, 1998
15. Глобальная спутниковая навигационная система ГЛОНАСС. - М.: МГУГиК (МИИГАиК), 1998
16. Дейт К. Руководство по реляционной СУБД DB2. – М.: Финансы и статистика, 1988 - 320 с.

17. Дмитриев В.И. К вопросу о государственной стратегии России в области CALS-технологий // Информационные технологии. - №5. - 1996 с.5- 9.
18. Дубинин В.Н., Зинкин С.А, Кулагин В.П. Система моделирования на основе сетей Петри // Автоматизация научных исследований, эргономического проектирования и испытаний сложных человеко-машинных систем, Всес. конф., ч.2. - Ленинград, 1983. - С.82-83.
19. Духно И.А., Чубуков Г.Л. Земельный порядок и ответственность М.: Институт защиты предпринимательства, 1998.
20. Жалковский Е.А., Халугин Е.И., Комаров А.И., Серпуховитин Б.И. Цифровая картография и геоинформатика. - М.: "Картоцентр-Геодэзиздат", 1999
21. Зинкин С.А., Кулагин В.П. Структуризация алгоритмов моделирования выполнения операций в процессоре базы данных. // Вопросы радиоэлектроники, сер. ЭВТ, вып.13, 1981, - С.78-87.
22. Зинкин С.А. , Кулагин В.П., О выполнении операций реляционной алгебры в имитационных системах . - // Вычислительная техника в автоматизированных системах контроля и управления:/ Межвуз. сб. научн. тр. Пенза: Пенз. политехн. ин-т, 1982, вып.12. - С.89-101.
23. Иванников А. Д. Решение систем многозначных логических уравнений при моделировании микропроцессорных систем // Электронное моделирование, т 9, 1987, №3 с 14-18
24. Иванников А. Д. Моделирование микропроцессорных систем - М.: Энергоатомиздат, 1990 -144с.
25. Иванников А.Д. Информационные технологии удаленного доступа и Концепция информатизации высшего образования./ в кн. ВУЗ и рынок, книга 4, Проблемы информатизации высшего образования России - М., "Пресс-сервис", 1994. - 216 с., с.73-91.
26. Иванников А.Д. Проблемы создания единой информационной среды высшей школы. Проблемы информатизации высшей школы (Метаинформация-координация-интеграция). 1995. N 1, с.4-1 - 4-2.
27. Иванников А.Д., Ижванов Ю.Л., Кулагин В.П. Перспективы использования WWW-технологии в высшей школе России. // Информационные технологии. 1996, №2, с. 24-29.
28. Иванников А.Д. , Кулагин В.П. Вопросы построения единого информационного пространства высшей школы России // Всероссийская научно-методическая конференция "Телематика'97", 19-23 мая, Санкт-Петербург. - 1997 г. - С.27-29



29. Информатика/под ред. Н.В. Макаровой. - М.: Финансы и статистика, 1997.
30. Информационные технологии в высшей школе. Официальные термины и понятия. ОСТ ВШ. Термины и определения 01.002-95. Дата введения 01.03.96- М.: Госкомвуз, 1996 -13с.
31. Карпик А.П., Середович А.В. Особенности формирования геоинформационных систем для целей кадастра // Вестник СГГА, 1999, Вып.3, с56-60.
32. Киенко Ю.П. Основы космического природоведения. - М.: Картоцентр-Геодезиздат, 1999 - 285 с.
33. Классификация автоматизированных систем и требования по защите информации. Защита от несанкционированного доступа к информации. - М.: Гостехкомиссия России, 1992.
34. Комов Н.В. и др. Земельные отношения и землеустройство в России. Москва 1995г
35. Кулагин В.П. Преобразование сложного запроса к процессору базы данных. // Вычислительная техника в автоматизированных системах контроля и управления: Межвуз. сб. научн. тр. Пенза: Пенз. политехн. ин-т, 1983, вып.13. - С.3-13.
36. Кулагин В.П. Формульное представление и правильность сетей Петри // Вычислительная техника в автоматизированных системах контроля и управления: Межвуз. сб. научн. тр. Пенза: Пенз. политехн. ин-т, 1984, вып.14. - С.34-39.
37. Кулагин В.П. Анализ вычислительных структур с использованием иерархических сетей // Вычислительная техника в автоматизированных системах контроля и управления: Межвуз. сб. научн. тр. Пенза: Пенз. политехн. ин-т, 1986, вып.16. - С.31-35
38. Кулагин В.П. Представление знаний в экспертной системе проектирования параллельных вычислительных структур // Матер. конф. "Математические и программные методы проектирования информационных и управляющих систем" / Пенза: ПДНТП, 1990,с.2
39. Кулагин В.П. Использование операций преобразования графов для синтеза вычислительных структур . - // Вопросы повышения эффективности систем управления технологическими процессами. / Сб. статей - Ереван, АРМТЕГ, 1991. - С.82-86.
40. Кулагин В.П. Анализ и синтез сложных структур как преобразование элементов линейного пространства // Вычислительная техника в автоматизированных системах контроля и управления: Межвуз. сб. научн. тр. Пенза: Пенз. политехн. ин-т, 1991, вып.21. - С.54-61.

41. Кулагин В.П. Алгебра сетевых моделей для описания параллельных вычислительных систем . - // Автоматизация и современные технологии. - 1993. - №2. - С.25-30
42. Кулагин В.П. Проблемы анализа и синтеза структур параллельных вычислительных систем . - // Информационные технологии. - 1997. - №1. - С.2-8
43. Кулагин В.П. Структуры сетей Петри . - // Информационные технологии. - 1997. - №4. -С.17-22
44. Кулагин В.П., Симонов А.В., Иванников А.Д. Концептуальные основы построения ГИС образовательных ресурсов России. ГИС для оптимизации природопользования в целях устойчивого развития территорий. Материалы Межд.конф., Барнаул, 1-4 июля 1998, изд-во Алтайского Гос.университета, 1998, с.18-19.
45. Кулагин В.П. Использование геоинформационных технологий в управлении отраслью на примере сферы образования России (доклад) // VI Межд. Форум «Информационные технологии и интеллектуальное обеспечение в здравоохранении и охране окружающей среды – 99»/Турция, Анталия, 3-10 октября, 1999 г. – С.74-77.
46. Кулагин В.П. Геоинформационные технологии – новый шаг на пути совершенствования системы управления образованием // IVВсероссийской конференции «Геоинформатика и образование» (Форум ГИС 2000). – М.: ГИС-Ассоциация, 2000 ,С.5-8
47. Кулагин В.П. Анализ научно-образовательных ресурсов России на основе геоинформационных технологий // Всероссийская научно-методическая конференция "Телематика'2000", Санкт-Петербург. - 1999 г. - С.3-10
48. Кулагин В.П., Константиновский Д.Л., Овсянников А.А., Симонов А.В., Цветков В.Я. Электронный атлас «Образовательные ресурсы России» // Тезисы докл. межд. конференции «ИОЛ-99» / С-Петербург, 2000 г. – С.268-270.
49. Кулагин В.П., Матчин В.Т. Конспект установочных лекций по дисциплине «Информатика» («Информсреда в образовании») // Учебное пособие /Под общ. Ред. В.А. Мордвинова /МГДТДиЮ, МИРЭА, М. – 2000.- 162 с.
50. Лазарев А.И., Савиных В.П. Достижения отечественной космонавтики в изучении окружающей среды. - СПб, Гидрометидат, 1996
51. Левин Б.А., Матвеев С.И., Цветков В.Я. Концепция создания геоинформационных систем железнодорожного транспорте /

Геодезия и геоинформатика в транспортном строительстве. - М.: МПС РФ, МГУПС (МИИТ), 2001. С. 3-7.

52. Лесных И.В., Середович В.-А., Парник А.-П., Горобцов А.В. Современные технологии инвентаризации земель для целей кадастра // Вестник СГГА, 1999, Вып.3, с53-56.
53. Лингер Р., Миллс Х., Уитт Б. Теория и практика структурного программирования. - М.: Мир, 1982
54. Лобанов А.Н., Журкин И.Г. Автоматизация фотограмметрических процессов. - М.: Недра, 1980
55. Лобанов В.С., Богатырь Б.Н., Иванников А.Д. Концепция информатизации высшего образования. Высшее образование в России, N 1, 1994, с.30-52.
56. Майкл Де Мере Географические информационные системы Основы. - М.: Дата+, 1999
57. Максимович Г.Ю., Романенко А.Г., Самойлюк О.Ф. Информационные системы./ Под общ. ред. К.И. Курбакова - М.: Из-во Рос. эконом. академ, 1999.
58. Максудова Л.Г., Савиных В.П., Цветков В.Я Интеграция наук об окружающем мире в геоинформатике // Исследование Земли из космоса.- №1. - 2000. с.46-50.
59. Максудова Л.Г. Цветков В.Я. Классификация информационных ресурсов // Геодезия и аэрофотосъемка, 2000, №4. с.140-149
60. Максудова Л.Г., Цветков В.Я. Информационное моделирование как фундаментальный метод познания. // Геодезия и аэрофотосъемка, 2001, №1. с.102-106
61. Максудова Л.Г., Цветков В.Я. Концепции создания единого информационного пространства геодезического образования в Российской Федерации // Геодезия и аэрофотосъемка, 2001, №4. с.147-152
62. Мейер М. Теория реляционных баз данных. – М.: Мир, 1987 - 608 с.
63. Мельников В. Защита информации в компьютерных системах. - М.: Финансы и статистика. Электронинформ, 1997
64. Месарович М., Мако Д., Такахага Н. Теория иерархических многоуровневых систем. - М.: Мир, 1973
65. Неумывакин Ю.К., Перский М.И. Автоматизированные методы геодезических измерений в землеустройстве. - М.: Недра, 1990.
66. Попов И. И., Храмцов П.Б. Мировые информационные ресурсы и сети. - М.: РЭА, 1999

67. Попович П.Р., Губинский А.И., Колесников Г.М., Савиных В.П. Системный анализ комплексов "Космонавт-техника". - М.: Машиностроение, 1994.
68. Постановление Правительства РФ от 16.01.95г. N 40 "Об организации работ по созданию геоинформационной системы для органов государственной власти"
69. Прорвич В.А. Основы экономической оценки городских земель. -М.: Дело, 1998.
70. Савиных В.П., Кучко А.П., Стеценко А.Ф. Аэрокосмическая фотосъемка. - М.: Картоцентр-Геодезиздат, 1997 - 378 с.
71. Савиных В.П., Малинников В.А., Сладкопевцев С. А. Цыпина Э.М. География из космоса. - М.: МГУГиК, 2000 - 234 с.
72. Савиных В.П., Цветков В.Я. Интеграция технологий ГИС и систем дистанционного зондирования Земли // Исследование Земли из космоса. №2.- 2000. с.1-4.
73. Савиных В.П., Цветков В.Я. Особенности интеграции геоинформационных технологий и технологий обработки данных дистанционного зондирования // Информационные технологии, 1999, №10. с. 36-40.
74. Савиных В.П., Цветков В.Я. Геоинформационные анализ данных дистанционного зондирования. М.: Картоцентр - Геодезиздат, 2001. - 224с.
75. Симонов А.В. Агроэкологическая картография. - Кишинев, 1991.
76. Сингор Р., Стегман М.О. Использование ODBC для доступа к базам данных. - М.: Бином, 1995
77. Соловьев Ю.А. Системы спутниковой навигации. М.: ЭКО-ТРЕНДЗ, 2000. - 268 с.
78. Тайли Эд Безопасность компьютера. - Минск, ООО Поппури, 1997
79. Тетерин Г.Н. Геодезия и построение картины окружающего мира // Геодезия и картография. 1997, №7, с. 48-51.
80. Тикунов В.С. Моделирование в картографии. - М.: Изд-во МГУ, 1997.
81. Тикунов В.С. Классификация в географии: ренессанс или увядание? - Москва- Смоленск.: Изд-во СГУ, 1997.
82. Тикунов В.С., Цапук Д.А. Устойчивое развитие территорий: Картографо-геоинформационное обеспечение - Москва- Смоленск.: Изд-во СГУ, 1999.
83. Тихонов А.Н. Единое информационное пространство высшей школы России: основные проблемы и направления развития // Информационные технологии. 1996, №2, с. 2-7.

84. Тихонов А.Н., Лобанов В.С., Иванников А.Д. Время информатизации // Высшее образование в России, N 2, 1996, с.30-33.
85. Тихонов А.Н., Иванников А.Д. Информатизация российского образования и общества в целом. Международное сотрудничество, 1997, N4, с.1-3.
86. Тихонов А.Н., Абрамешин А.Е., Воронина Т.П., Молчанова О.П. Иванников А.Д. Управление современным образованием: социальные и экономические аспекты / под ред. Тихонова А.Н. - М.: Вита-Пресс, 1998. - 256 с.
87. Тюфлин Ю.С. Космическая фотограмметрия при изучении планет и спутников. - М.: Недра, 1986.
88. Уемов А.И. Системный подход и общая теория систем. - М. Мысль, 1978
89. Урмаев М.С. Космическая фотограмметрия. - М.: Недра 1989
90. Хаксольд В. Введение в городские географические информационные системы. - Издательство Оксфордского университета, 1991 г.
91. Храмцов П.Б. Лабиринт Интернет. Практическое руководство. - М. Электроинформ, 1996
92. Цветков В.Я. Информатизация: Создание современных информационных технологий. Часть 1. Структуры данных и технические средства. - М., ГКНТ, ВНИЦентр, 1990.
93. Цветков В.Я. Разработка проблемно ориентированных систем управления. -М.: ГКНТ, ВНИЦентр, 1991.
94. Цветков В.Я. Методы и системы обработки и представления видеонформации. - М.:ГКНТ, ВНИЦентр, 1991.
95. Цветков В.Я. Проектирование структур данных и базы данных / Учебное пособие - МГУГиК - М 1997.- 90 с.
96. Цветков В.Я. Особенности развития информационных стандартов в области новых информационных технологий // Информационные технологии, №8, 1998 с 2-7.
97. Цветков В.Я. Особенности стандартизации информационных средств и продуктов // Геодезия и аэрофотосъемка. 1998- N2 с 118 - 124.
98. Цветков В.Я. Геоинформационные системы и технологии. - М.: Финансы и статистика, 1998.
99. Цветков В.Я. Геоинформационное моделирование // Информационные технологии, 1999, №3. с. 23- 27.
100. Цветков В.Я. Методы прогнозирования в геоинформационных системах // Информатика - машиностроение, 1999, № 1. с. 11-13

101. Цветков В.Я. Особенности защиты информации в геоинформационных технологиях // Информатика - машиностроение, 1999, № 4. с. 44-46.
102. Цветков В.Я., Мощиль В.И. Принятие решений в условиях риска. М., Геодезия и аэрофотосъемка, 1999, №4, с.92-97.
103. Цветков В.Я. Автоматизированные системы обработки экономической информации./ ЭИ - М.: МГУГиК, 1999 - 106 с.
104. Цветков В.Я. Проблемы интеллектуализации геоинформационных систем // Информатика - машиностроение, 1999, № 5-6. с. 30-31.
- 105.. Цветков В.Я. Геоинформационные системы и технологии - М.: МГУГиК, 2000 -69 с.
106. Цветков В.Я. Геомаркетинг. - М.: Машиностроение, 2000 - 64 с.
107. Цветков В.Я. Стандартизация информационных программных средств и программных продуктов. - М.: МГУГиК, 2000. - 86 с.
108. Цветков В.Я. Методологические основы тестирования информационных и геоинформационных систем // Геодезия и аэрофотосъемка, 2000, №3. с.93-102.
109. Цветков В.Я. Цифровые карты и цифровые модели // Геодезия и аэрофотосъемка, 2000, №2. с.147-155
110. Цветков В.Я. Создание интегрированной информационной основы ГИС // Геодезия и аэрофотосъемка, 2000, №4. с.150-154
111. Цветков В.Я., Жукова О.С. Поддержка принятия решений в геоинформационных системах // Машиностроитель. - №1- 2000. - с. 28- 30.
112. Цветков В.Я. Принятие решений в условиях неопределенности // Техника машиностроения, 2000, №6, с. 17- 21
113. Цветков В.Я. ГИС как система визуальной обработки информации // Геодезия и аэрофотосъемка, 2000, №2. с.143-147
114. Цветков В.Я., Максудова Л.Г. Особенности подготовки специалистов в области информатики // Машиностроитель. - №5- 2000. - с. 34- 35
115. Цветков В.Я. Защита информации в системах обработки данных и управления. - М.: Миннауки, ВНИИЦ, 2000.
116. Цветков В.Я. Применение геоинформационных технологий для поддержки принятия решений // Геодезия и аэрофотосъемка, 2001, №4. с.128-138
117. Цветков В.Я. Информационная безопасность и геоинформационные технологии // Информационные технологии, №7, 2000 с 25-32.

118. Цветков В.Я., Журкин И.Г. Особенности защиты информации в геоинформационных технологиях/ в кн. "220 лет геодезическому образованию в России": Сборник трудов Международной научно-технической конференции, посвященной 220-летию со дня основания Московского государственного университета геодезии и картографии (Москва 24-29 мая 1999 г.) в 2-х частях. Часть II - М.: МГУГиК, 2000. С. 156- 163.
119. Цветков В.Я. Стандартизация информационных программных средств и программных продуктов - М.: МГУГиК, 2000 116 с
120. Цветков В.Я Информационная безопасность и геоинформатика // Геодезия и аэрофотосъемка, 2001, №1. с.106-121
121. Цветков В.Я. Автоматизированные земельные информационные системы - М.: Минпромнауки, 2001- 68 с.
122. Цветков В.Я. Основы системного подхода при анализе информационных систем // Геодезия и аэрофотосъемка, 2001, № 4. с. 111-115
123. Цветков В.Я. Методы и системы поддержки принятия решений в управлении. - М.: Минпромнауки, 2001- 76 с.
124. Цикритзис Д., Лоховски Ф. Модели данных. - М.: Финансы и статистика, 1986.
125. Шатихин Л.Г. Структурные матрицы и их применение для исследования систем. - М.: Машиностроение, 1974
126. Шенон Р. Имитационное моделирование - искусство и наука. - М. Мир, 1978
127. Bozzelti The Technological Evolution of ICT and Standards // European Information Technology Observatory - EITO 97 (Part One), 1997. p. 88 -95
128. Snyder J.P., Voxland P.M. An Album of Map Projection. U.S. Geological Survey professional paper 1453, Washington, 1989
129. Department of Defense Trusted Computer System Evaluation Criteria -- DoD 5200.28-STD, 1993
130. George B. Karte, P.E. The GIS Book/- USA, JnWordPress, 1997.
131. Information Technology Security Evaluation Criteria (ITSEC). Harmonised Criteria of France — Germany — the Netherlands — the United Kingdom -- Department of Trade and Industry, London, 1991
132. A.A.Alekseyev, A. D. Ivannikov, S.V.Mischenko, N.S.Popov, V.Y.Podolskyi Project «Encyclopaedia of Computer Modelling» (ENCOM) in the System of Distance Education of Higher Education Institution - Proc. of the Second Int. Conf. On Distance Education in

Russia: Open and Distance Learning as a Development Strategy,  
1996, Moscow, vol. 2, pp.386-388

133. A. Simonov, E. Kiknadze, A. Ivannikov GIS Education in Russia  
Higher School: Concept and Actions. First Int. Conf. in GIS  
Education, Marriott at Eagle Crest Ypsilanti, Michigan, USA,  
Oct., 29-Nov., 1, 1998, pp/28-29.



## Содержание

Введение.	3
<i>Общая геоинформатика</i>	
Информационные ресурсы и модели	10
1.1. От информации к информационным ресурсам	10
1.2. Модели информационных ресурсов	15
1.3. Информационное моделирование как метод познания	19
1.4. Семиотический подход к организации информационных моделей	23
2. Основные виды и характеристики моделей	24
2.1. Сильно и слабо типизированные модели	24
2.2. Статические и динамические модели	25
2.3. Аналоговые и дискретные модели	26
2.4. Масштаб действия модели	26
2.5. Жизненный цикл модели	27
2.6. Формы представления моделей данных	28
3. Базовые модели данных, используемые в геоинформатике	30
3.1. Иерархическая модель	30
3.2. Квадратомическое дерево	34
3.3. Реляционная модель	37
3.4. Инфологическая модель данных "Сущность-связь"	39
3.5. Сетевые модели	44
4. Сбор информации	48
4.1. Первичная обработка информации	50
4.2. Кодирование информации	50
5. Интеграция данных в геоинформатике	55
6. Теоретические основы обработки данных в геоинформатике	58
6.1. Информационные процессы	58
6.2. Описания систем через множество процессов.	61
6.2.1. Сети Петри	61
6.2.2. Механизм агентов	68
6.2.3. Предикаты над процессами	75
6.3. Вычислимость в геоинформатике	76
6.4. Фракталы как метод описания пространственных данных	78
6.5. Обработки графической информации	83
6.5.1. Графические изображения	83

6.5.2. Графические редакторы	87
6.6. Объектно-ориентированный подход в геоинформатике	88
7. Хранение данных	94
7.1. Проектирование реляционных баз данных	94
7.2. Универсальное отношение	95
7.3. Проблемы, требующие разбиения таблицы	95
7.4. Нормализации как метод построения нормальных форм	96
7.5. Функциональные и многозначные зависимости	97
7.6. Нормальные формы	98
7.7. Процедура нормализации	99
8. Основы системного подхода при анализе и проектировании геоинформационных систем	101
8.1. Системный анализ	101
8.2. Системное проектирование в геоинформатике	107
<i>Прикладная геоинформатика</i>	
9. Сбор пространственно-временных данных в геоинформатике	113
9.1. Исходные и унифицированные данные	113
9.2. Основные характеристики пространственных объектов	115
9.3. Форма и формат данных	118
9.4. Основные технологии сбора данных	121
9.5. Системы сбора полевой информации	124
96. Система глобального позиционирования GPS	124
9.7. Сбор картографических данных	126
9.8. Фотограмметрические методы сбора данных	129
9.9. Применение методов сбора и обработки ДДЗ в геоинформатике	132
9.9.1. Многозональная съемка	133
9.9.2. Инфракрасная съемка	133
9.9.3. Радиолокационная съемка	134
10. Координатные и атрибутивные данные и модели.	136
10.1. Координатные модели	136
10.1.1. Точечные модели	138
10.1.2. Линейные модели	139
10.1.3. Площадные модели	139
10.2 Атрибутивные модели	140
10.3. Точность координатных и атрибутивных данных	141
10.4. Векторные и растровые модели	142

10.4.1. Векторные модели	143
10.4.2. Топологические модели и характеристики	144
10.5. Растровые модели	148
10.5.1. Характеристики растровых моделей.	149
11. Геоинформационные системы (ГИС)	155
11.1. Развитие и определение ГИС	155
11.2. Особенности ГИС	159
12. Цифровое моделирование и цифровые модели	163
12.1. Основные понятия	163
12.2. Математическая основа цифровых моделей	165
12.3. Геодезическая система координат	167
12.4. Системы координат	168
12.5. Картографические проекции	170
12.6. Цифровые карты	177
12.7. Характеристики цифровых моделей	178
12.8. Виды моделирования	181
12.9. Цифровые модели рельефа	183
12.10. Построение цифровых моделей рельефа на основе метода Вороного-Делоне	183
12.10.1. Историческая справка	183
12.10.2. Метод Вороного-Делоне	185
13. Геоинформационное моделирование	188
13.1. Формализация технологических процессов	189
13.2. Комбинирование выбранных объектов	190
13.3. Преобразование атрибутов комбинируемых объектов	191
13.4. Геоинформационное моделирование с использованием геогрупп	193
13.5. Построение новых графических объектов на основе слияния атрибутивных данных	194
13.6. Буферизация	194
13.7. Геокодирование	195
13.8. Классификация	197
14. Основные технологические этапы обработки данных в ГИС	198
14.1. Основные функции интерфейса ГИС	198
14.2. Подготовительный этап	200
14.3. Применение файлов в формате DXF для создания картографической основы	201
14.4. Работа с растровыми данными	201
14.4.1. Ввод растрового изображения	201

14.4.2. Векторизация растрового изображения	203
14.5. Редактирование графических данных	204
14.6. Создание точечных объектов	206
14.7. Вывод информации	207
14.8. Методы обработки данных дистанционного зондирования	209
15. ГИС как система управления	213
16. Земельный кадастр	221
16.1. Кадастр как система налогообложения	221
16.2. Исторический обзор развития кадастра.	221
16.3. Развитие Кадастра в России	224
16.4. Современное состояние кадастра за рубежом	225
16.5. Государственный земельный кадастр	226
17. Земельные информационные системы	227
17.1. Применение ЗИС на операционном уровне управления	228
17.2. Применение ЗИС на тактическом уровне управления	230
17.3. Применение ЗИС на стратегическом уровне управления	230
18. Геоинформационное прогнозирование	231
18.1. Анализ, моделирование вторичных (унифицированных) данных	232
18.2. Методы прогиозирования	233
18.3. Прогнозные модели и их характеристики	234
18.4. Верификация прогноза	235
18.5. Выбор метода прогнозирования	235
19. Геомаркетинг	236
19.1. Виды геомаркетинга	240
19.1.1. Геомаркетинг мест	240
19.1.2. Природоресурсный геомаркетинг	241
19.1.3. Дополнителные виды геомаркетинга	243
19.1.4. Политический геомаркетинг	244
<i>Специальная геоинформатика</i>	
20. Выбор аппаратно-программной платформы	246
21. Классификация компьютеров, применяемых в геоинформационных технологиях	248
21.1. Персональные компьютеры и рабочие станции	248
21.2. Серверы	250
21.3. Мейнфреймы	252
21.4. Кластерные архитектуры	253

22. Кодирование и элементы теории информации.	257
22.1. Постановка вопроса.	257
22.2. Коды и кодирование	258
22.3. Коды постоянной длины	260
22.4. Коды переменной длины	261
22.5. Последовательное и параллельное кодирование	262
22.6. Оптимальность кодов и разрешающая информация	263
22.7. Надежность передачи сообщений	264
22.8. Надежность кодов	265
22.9. Надежность передачи кодов	266
23. Представление пространственных объектов в виде множеств	270
23.1. Вычислительная структура множеств с доступом по ключу	271
23.2. Метод хэширования	271
24. Информационная безопасность	276
24.1. Геоинформатика и информационная безопасность	276
24.2. Система безопасности	277
24.3. Показатели защищенности средств вычислительной техники от несанкционированного доступа	278
24.4. Классы защищенности автоматизированных систем	280
24.5. Каналы доступа к информационным системам	284
24.6. Группы методов несанкционированного доступа	285
24.7. Информационные угрозы	286
24.8. Предотвращение угроз	286
24.9. Политика безопасности и уровни ее реализации	287
24.10. Методы защиты от компьютерных вирусов	289
24.11. Виды возможных последствий НСД к информации ГИС	290
24.12. Функции системы информационной безопасности в геоинформатике	291
24.13. Оценки характеристики информационной безопасности ГИС	293
24.14. Защита информационных ресурсов в геоинформатике	294
24.15. Уровни информационной защиты в ГИС	295
24.16. Тактика и стратегия защиты информации в геоинформатике	296

24.17. Информационная безопасность и Интернет	296
24.18. Условия безопасности передачи данных по сети	297
25. Стандартизация информационных продуктов и систем в геоинформатике	299
25.1. Роль стандартизации в сфере геоинформатики	299
25.2. Два типа стандартов в области новых информационных технологий	300
25.3. Концепции разработки стандартов	301
25.4. Стандартизация информационных продуктов и систем в международных организациях	302
25.5. Особенности стандартизации инструментальных программных средств и программных продуктов	305
25.6. Основные положения стандартизации	307
25.7. Объект стандартизации	308
25.8. Основные цели и задачи стандартизации	308
25.9. Категории стандартизации	309
25.10. Виды стандартов	309
25.11. Жизненный цикл продукции	310
25.12. Система качества	311
25.13. Качество информации	313
26. Сертификация информационных средств и программных продуктов	314
26.1. Цели сертификации ИПС	315
26.2. Заявка на сертификацию	316
26.3. Условия сертификации	316
26.4. Процесс сертификации	318
26.5. Результаты сертификации	318
26.6. Оплата работ по сертификации	319
26.7. Срок действия сертификата	319
Список сокращений	320
Терминологический словарь	322
Литература	334

*Научное издание*

**ИВАННИКОВ Александр Дмитриевич,  
КУЛАГИН Владимир Петрович,  
ТИХОНОВ Александр Николаевич,  
ЦВЕТКОВ Виктор Яковлевич**

## **ГЕОИНФОРМАТИКА**

**Напечатано с готового оригинал-макета**

**Издательство ООО "МАКС Пресс"  
Лицензия ИД N 00510 от 01.12.99 г.**

**Подписано в печать 04.10.2001. Формат 60×84<sup>1</sup>/16.**

**Печать офсетная. Бумага офсетная № 1.**

**Усл. печ. л. 22,0. Тираж 1000 экз. Заказ 7792.**

**"МАКС Пресс". 107066, г. Москва, Елоховский пр., д. 3, стр. 2.**

**Тел. 939-38-90, 939-38-91, 928-10-42. Тел./Факс 939-38-91.**

**Отпечатано в Производственно-издательском комбинате ВИНТИ,**

**140010, г. Люберцы, Московской обл., Октябрьский пр-т, 403.**

**Тел. 554-21-86.**

А.Д. Иванников, В.П. Кулагин,  
А.Н. Тихонов, В.Я. Цветков

ГЕОИНФОРМАТИКА