

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ УКРАИНЫ
ХАРЬКОВСКАЯ НАЦИОНАЛЬНАЯ АКАДЕМИЯ ГОРОДСКОГО ХОЗЯЙСТВА

В. Д. Шипулин

ОСНОВНЫЕ ПРИНЦИПЫ
ГЕОИНФОРМАЦИОННЫХ СИСТЕМ

Учебное пособие

Харьков
ХНАГХ
2010

УДК 004:910:528(075)

ББК 26+32.965Я7

Ш84

Рецензент:

Палеха Ю. Н., доктор географических наук, заместитель директора по науке, директор Базового центра ГИС Киевского государственного института проектирования городов

Рекомендовано к печати решением Ученого совета Харьковской национальной академии городского хозяйства. Протокол № 4 от 27.12.2009 г.

Шипулин В. Д.

Ш84 Основные принципы геоинформационных систем: учебн. пособие / Шипулин В. Д.; Харьков. нац. акад. гор. хоз-ва. – Х.: ХНАГХ, 2010. – 337 с.

В пособии представлены основные концепции и принципы построения и функционирования географических информационных систем (ГИС). Рассмотрены компьютерные модели географических объектов, вопросы сбора и подготовки географических данных, организации данных в геоинформационных системах, основы геопространственного анализа, которые составляют содержание учебной дисциплины "Основы ГИС".

Для студентов специальности "Геоинформационные системы и технологии", а также студентов и специалистов в области компьютерных наук, управления территориями, земельными ресурсами и недвижимостью, коммунальным хозяйством, транспортной инфраструктурой, энергетики, экологии, в областях разработки, создания или исследования сложных пространственных комплексов.

УДК 004:910:528(075)

ББК 26+32.965Я7

© Шипулин В. Д., 2010

© Харьков, ХНАГХ

СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ	12
1 ВВЕДЕНИЕ В ГЕОГРАФИЧЕСКИЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ	16
1.1 Концепция систем	17
1.1.1 Сложность и взаимозависимость частей нашего мира	17
1.1.2 Исторический аспект	18
1.1.3 Определения понятия "Система"	19
1.1.4 Системные исследования	20
1.1.5 Элементы систем	20
1.1.6 Системные связи	21
1.1.7 Понятие "Структура"	22
1.1.8 Системный эффект	23
1.1.9 Системные принципы	23
1.1.10 Формы и способы описания систем	24
1.1.11 Классификация систем	27
1.1.12 Контрольные вопросы и задания для самостоятельной работы	28
1.2 Общая характеристика географических информационных систем	29
1.2.1 Область ГИС	29
1.2.2 Определение ГИС	31
1.2.3 Общая характеристика компонентов ГИС	33
1.2.3.1 Аппаратные средства ГИС	33
1.2.3.2 Программное обеспечение ГИС	34
1.2.3.3 Географические данные	38
1.2.3.4 Регламент	40
1.2.3.5 Пользователи	40
1.2.4 Общая характеристика отличительных функций ГИС	40
1.2.4.1 Интегрирование	41
1.2.4.2 Анализ	42
1.2.4.3 Визуализация	44
1.2.5 ГИС как пересечение областей научных знаний	45
1.2.6 Примеры применения ГИС	46
1.2.7 Значение ГИС для управления городом	49

1.2.8	Историческая справка о геоинформационных системах	49
1.2.9	Контрольные вопросы и задания для самостоятельной работы	55
1.3	ГИС-парадигма	56
1.3.1	Что есть Географическая информация, географические данные?	56
1.3.2	Что есть Геоинформационные технологии?	57
1.3.3	Что есть Географическая информационная система?	58
1.3.4	Что есть Геоинформационная наука?	59
1.3.5	Что есть Геоинформационное образование?	61
1.3.6	Что есть Геоинформатика?	63
1.3.7	Контрольные вопросы и задания для самостоятельной работы	64
2	КОМПЬЮТЕРНЫЕ МОДЕЛИ ГЕОГРАФИЧЕСКИХ ОБЪЕКТОВ	65
2.1	Географические объекты	66
2.1.1	Пространство и время	66
2.1.2	Определение географических объектов	67
2.1.3	Виды географических объектов	67
2.1.4	Способы локализации географических объектов	68
2.1.5	Контрольные вопросы и задания для самостоятельной работы	70
2.2	Географические данные	71
2.2.1	Понятия "информация" и "данные"	71
2.2.2	Географическая информация и географические данные	73
2.2.3	Атрибуты пространственных объектов	75
2.2.4	Виды компьютерных моделей географических объектов	76
2.2.5	Контрольные вопросы и задания для самостоятельной работы	77
2.3	Векторные модели географических объектов	78
2.3.1	Базовые понятия картографического представления географических объектов	78
2.3.2	Простые векторные модели географических объектов	79
2.3.2.1	Простые векторные модели единичных географических объектов	79

2.3.2.2	Простые векторные модели множества географических объектов	80
2.3.3	Топологические векторные модели географических объектов	82
2.3.3.1	Понятие "Граф"	82
2.3.3.2	Понятие "Пространственные отношения"	83
2.3.3.3	Понятие "Топология"	84
2.3.3.4	Топологическое представление области	85
2.3.3.5	Топологическое представление смежности	86
2.3.3.6	Топологическое представление связности	88
2.3.3.7	Сравнение векторных моделей географических объектов	90
2.3.4	Форматы векторных данных	90
2.3.5	Резюме представления географических объектов векторными моделями	92
2.3.6	Контрольные вопросы и задания для самостоятельной работы	93
2.4	Растровые модели географических объектов	94
2.4.1	Концепция растровых моделей географических объектов	94
2.4.1.1	Определение и виды растровых моделей	94
2.4.1.2	Источники растровых данных	95
2.4.1.3	Матрица ячеек	96
2.4.1.4	Представление дискретных географических объектов	96
2.4.1.5	Представление непрерывных явлений регулярными моделями	98
2.4.1.6	Недостатки и преимущества растровых моделей	101
2.4.2	Характеристики растровых моделей	101
2.4.2.1	Разрешение	102
2.4.2.2	Геометрия растров	102
2.4.2.3	Координаты ячеек	103
2.4.2.4	Экстент региона	104
2.4.2.5	Топология ячеек растров	105
2.4.2.6	Значение ячеек растра	105
2.4.2.7	Кодирование значений	106

2.4.2.8	Цветовые модели	106
2.4.2.9	Зонирование	108
2.4.2.10	Таблица атрибутов растра	109
2.4.2.11	Растровые слои	109
2.4.3	Хранение растровых данных	110
2.4.4	Сжатие растровых данных	112
2.4.4.1	Метод группового кодирования	112
2.4.4.2	Методы, основанные на порядке сканирования	114
2.4.4.3	Квадратомическое дерево	115
2.4.5	Форматы растровых данных	117
2.4.6	Файл геопривязки растровых данных	119
2.4.7	Резюме представления географических объектов растровыми моделями	120
2.4.8	Контрольные вопросы и задания для самостоятельной работы	122
2.5	Триангуляционные модели географических объектов	123
2.5.1	Определение модели TIN	123
2.5.2	Свойства модели TIN	124
2.5.3	Триангуляция Делоне	125
2.5.4	Топология в TIN	126
2.5.5	Этапы создания модели TIN	128
2.5.6	Уравнение нормали к грани треугольника	129
2.5.7	Средства модели TIN для отображения поверхности	131
2.5.7.1	Экспозиция склона	131
2.5.7.2	Крутизна грани	131
2.5.7.3	Затенение граней	132
2.5.7.4	Диапазоны высот	133
2.5.8	Резюме представления географических объектов триангуляционными моделями	133
2.5.9	Контрольные вопросы и задания для самостоятельной работы	135
3	СБОР И ПОДГОТОВКА ГЕОГРАФИЧЕСКИХ ДАННЫХ	136
3.1	Сбор и предварительная обработка географических данных	137
3.1.1	Источники географических данных	137
3.1.2	Характеристики данных	148

3.1.3	Системный подход к предварительной обработке исходных данных	150
3.1.4	Контрольные вопросы и задания для самостоятельной работы	156
3.2	Классификация	157
3.2.1	Общие положения классификаций	157
3.2.1.1	Классификация — средство упорядочения знаний	157
3.2.1.2	Сложность классификации	159
3.2.1.3	Подходы классификации	159
3.2.1.4	Значение классификации для геоинформационных систем	161
3.2.2	Базовые понятия единой системы классификации технико-экономической информации	161
3.2.2.1	Термины и определения классификации	162
3.2.2.2	Термины и определения кодирования	163
3.2.2.3	Организационные аспекты классификации	164
3.2.2.4	Особенности создания системы классификации	165
3.2.2.5	Единая система классификации и кодирования технико-экономической информации	171
3.2.3	Классификации в геоинформационных системах	173
3.2.3.1	Объекты классификации в геоинформационных системах	173
3.2.3.2	Примеры классификаций	175
3.2.4	Контрольные вопросы и задания для самостоятельной работы	184
3.3	Системы координат и картографические проекции	185
3.3.1	Значение координат в геоинформационных системах	185
3.3.2	Земные сферы и сфероиды	186
3.3.3	Геодезические даты	189
3.3.4	Системы координат	191
3.3.4.1	Мировая геодезическая система координат WGS-84	191
3.3.4.2	Системы криволинейных координат	193

3.3.4.3	Проецированные системы координат	196
3.3.5	Картографические проекции	196
3.3.5.1	Что представляет собой картографическая проекция?	196
3.3.5.2	Классификация проекций по типу искажений	198
3.3.5.3	Классификация проекций по типу поверхности проецирования	199
3.3.5.4	Другие проекции	205
3.3.5.5	Параметры картографических проекций	206
3.3.5.6	Примеры картографических проекций	207
3.3.6	Географические преобразования	213
3.3.6.1	Метод с тремя параметрами	213
3.3.6.2	Метод с семью параметрами	214
3.3.6.3	Метод Молоденского	215
3.3.7	Контрольные вопросы и задания для самостоятельной работы	216
4	ОРГАНИЗАЦИЯ ДАННЫХ В ГЕОИНФОРМАЦИОННЫХ СИСТЕМАХ	218
4.1	Понятие системной организация данных	219
4.1.1	Определение, значение и задачи системной организация данных	219
4.1.2	Уровни организации данных в ГИС	221
4.1.3	Принципы организации данных в ГИС	226
4.1.4	Виды моделей организации данных	228
4.1.5	Контрольные вопросы и задания для самостоятельной работы	229
4.2	Геореляционная модель данных	230
4.2.1	Сущность геореляционной модели данных	230
4.2.2	Слои пространственных данных - вертикальная организация данных	231
4.2.3	Пространственная индексация - горизонтальная организация данных	236
4.2.4	Модель данных "Шейпфайл"	237
4.2.5	Модель данных "Покрытие"	239
4.2.6	Преимущества и недостатки геореляционной модели данных	242
4.2.7	Контрольные вопросы и задания для самостоятельной работы	243
4.3	Объектно-ориентированная модель данных	244

4.3.1	Основные положения объектно-ориентированной методологии	244
4.3.1.1	Понятие объекта	244
4.3.1.2	Понятие класса	245
4.3.1.3	Понятие отношения	246
4.3.1.4	Принципы объектной модели	247
4.3.2	Общая характеристика объектно-ориентированной модели данных "База геоданных"	248
4.3.2.1	Определение базы геоданных	248
4.3.2.2	Объектно-ориентированная векторная модель данных	249
4.3.2.3	Преимущества базы геоданных	249
4.3.3	Средства придания интеллектуальных свойств пространственным объектам	250
4.3.3.1	Топология	251
4.3.3.2	Подтипы	253
4.3.3.3	Домены	254
4.3.3.4	Отношения и классы отношений	255
4.3.4	Элементы БГД	258
4.3.4.1	Таблицы	259
4.3.4.2	Расширения таблиц	260
4.3.4.3	Классы пространственных объектов	260
4.3.4.4	Расширения классов пространственных объектов	261
4.3.4.5	Наборы растровых данных	262
4.3.4.6	Расширения растров	263
4.3.5	Типы баз геоданных	264
4.3.6	Преимущества и недостатки объектно-ориентированной модели данных	265
4.3.7	Контрольные вопросы и задания для самостоятельной работы	266
5	ОСНОВЫ ГЕОПРОСТРАНСТВЕННОГО АНАЛИЗА	267
5.1	Общая характеристика геопространственного анализа	268
5.1.1	Определение геопространственного анализа	268
5.1.2	Задачи геопространственного анализа	269
5.1.3	Методология геопространственного анализа	270
5.1.4	Классификации аналитических средств ГИС	274
5.1.5	Контрольные вопросы и задания для самостоятельной работы	275

5.2	Функции измерений	276
5.2.1	Измерения на векторных данных	276
5.2.2	Измерения на растровых данных	281
5.2.3	Контрольные вопросы и задания для самостоятельной работы	282
5.3	Функции выбора данных	283
5.3.1	Интерактивный пространственный выбор данных	283
5.3.2	Пространственный выбор по атрибутивным условиям	283
5.3.3	Пространственный выбор на основании топологических отношений	284
5.3.4	Контрольные вопросы и задания для самостоятельной работы	285
5.4	Функции классификации	286
5.4.1	Цели классификации по атрибутам	286
5.4.2	Методы автоматизированной классификации по атрибутам	287
5.4.3	Контрольные вопросы и задания для самостоятельной работы	289
5.5	Оверлейные функции	290
5.5.1	Определение и общая характеристика оверлейных функций	290
5.5.2	Булева алгебра в топологическом наложении	292
5.5.3	Векторные оверлейные операторы	293
5.5.3.1	Классификация векторных оверлейных операций	293
5.5.3.2	Алгоритмы векторных оверлейных операций	296
5.5.3.3	Базовые оверлейные операции векторных моделей	299
5.5.4	Растровые оверлейные операторы	302
5.5.5	Контрольные вопросы и задания для самостоятельной работы	305
5.6	Функции окрестности	306
5.6.1	Определение окрестности	306
5.6.2	Операции окрестности в векторных моделях	307
5.6.2.1	Генерирование буферных зон	307
5.6.2.2	Генерирование полигонов Тиссена	312
5.6.3	Операции окрестности в растровых моделях	313
5.6.3.1	Функции фокальной статистики	313

5.6.3.2	Функции распространения	316
5.6.3.3	Функции распределения	320
5.6.4	Контрольные вопросы и задания для самостоятельной работы	321
5.7	Функции связности	322
5.7.1	Определение и характеристика сети	322
5.7.2	Нахождение лучшего пути	324
5.7.3	Разделение сети	325
5.7.4	Контрольные вопросы и задания для самостоятельной работы	329
	ПОСЛЕСЛОВИЕ	330
	БИБЛИОГРАФИЧЕСКОЕ ОПИСАНИЕ	332
	Использованная литература	332
	Рекомендуемая литература для самостоятельного изучения	336

ПРЕДИСЛОВИЕ

Интерес к географическим информационным системам (ГИС), геоинформационным технологиям и геоинформационной науке имеет чрезвычайный рост в последние годы благодаря полученным на их основе эффективным решениям во многих областях человеческой деятельности. С их помощью решаются локальные, региональные и глобальные задачи устойчивого развития территорий, использования природных ресурсов, охраны окружающей среды, обеспечения общественной безопасности. Современные геоинформационные системы расширяют методы исследования нашего мира, предоставляя цифровые инструменты для организации и оперирования пространственными данными, моделирования происходящих в пространстве процессов, визуализации этих данных, моделей и процессов с помощью развитых компьютерных средств, специализированных инструментов обработки и анализа геоданных. Исследователи от разнообразных академических дисциплин используют пространственное мышление и инструменты ГИС для развития явных пространственных моделей на разных масштабных уровнях. Они используют геоинформационные системы как средство понимания мира, описывая и объясняя отношения человечества к реальному миру. Широкий общественный интерес к этой теме подогревается популярными приложениями подобным картам Google, персональным спутниковым навигационным устройствам, Интернет картографическим сервисом MapQuest, и т.п. Пользователи ГИС все в большей мере анализируют сущности и явления реального мира, интегрируя информацию из разных источников с цифровыми картами веб-сервисов, таких как, например, Google Earth и MS Virtual Earth.

В ответ на этот вызов научно-технического прогресса развивается область ГИС-образования. В вузах для ряда специальностей введены отдельные курсы по геоинформационным системам или геоинформатике, а для специальностей по информационным технологиям - соответствующие специализации. В 2001 году решением Кабинета Министров Украины в перечень специальностей введена специальность "Геоинформационные системы и технологии".

В мире выходит огромное количество статей, материалов конференций, книг по многочисленным темам геоинформационных систем и технологий, которые предназначены в большей мере для специалистов ГИС или конечных пользователей популярного

программного обеспечения ГИС, чем для начинающих. Поэтому есть необходимость в этом "океане" знаний выделить ключевые концепции и принципы ГИС и донести их разнообразной аудитории студентов в форме учебной литературы.

Предлагаемое вниманию читателя пособие охватывает учебную дисциплину "Основы ГИС" специальности "Геоинформационные системы и технологии". Пособие призвано способствовать созданию теоретической базы изучения ГИС. В пособии представлены основные концепции, положения, понятия, принципы построения и функционирования географических информационных систем. Пособие не является практикумом по освоению программного обеспечения ГИС. Пособие построено из расчета того, что в соответствии с учебным планом специальности: 1) студенты уже имеют определенную подготовку по компьютерным дисциплинам и дисциплинам, изучающим земное пространство, 2) после изучения основ ГИС студенты получают знания по развитым технологиям ГИС, выполнению геоинформационного анализа, дистанционному зондированию Земли, приложениям ГИС в различных областях.

Сегодня в мире разработано огромное количество программных продуктов ГИС. В рамках пособия даже краткий обзор программных продуктов ГИС в виде перечня характеристик или функций не может создать адекватную картину программного вооружения ГИС. Темпы создания нового программного обеспечения ГИС или модернизации созданного столь высоки, что существует опасность освещения в пособии устаревших решений. Настоящее пособие построено в большей мере в контексте решений Института исследования систем окружающей среды (Environmental Systems Research Institute - ESRI), который создал наиболее развитое и распространенное в мире программное обеспечение ГИС.

Пособие состоит из пяти частей, каждая из которых отображает определенную предметную область знаний о географических информационных системах. Эти части являются модулями - однородными по тематике структурными единицами для изучения учебной дисциплины. Пособие построено таким образом, чтобы в логической последовательности вводить читателя в мир ГИС.

Первая часть "Введение в географические информационные системы" начинается с рассмотрения фундамента знаний о географических информационных системах - понятийной базы теории систем, общих закономерностей развития и функционирования систем. В разделе представлена общая характеристика географических информационных систем, дано определение ГИС, рассмотрены общая

характеристика компонентов ГИС и отличительных функций ГИС. Как ГИС-парадигма представлена единая система базовых понятий и определений на основе устоявшихся взглядов, которая является каркасом представления предметной области геоинформационных систем и технологий.

Вторая часть "Компьютерные модели географических объектов" содержит определение географических объектов и их видов, определение географической информации и географических данных. В отдельных разделах этой части рассматриваются принципы построения и характеристики базовых компьютерных моделей географических объектов - векторных, растровых, триангуляционных, которые проверены временем и реализованы в подавляющем большинстве современных программных продуктов ГИС.

Третья часть "Сбор и подготовка географических данных" начинается с рассмотрения источников географических данных, важнейших характеристик данных, системного подхода к предварительной обработке исходных данных. Большое внимание уделяется проблемам классификаций, которые лежат в основе эффективной работы геоинформационных систем. Значительное место отводится рассмотрению систем координат реального земного пространства и картографическим проекциям как фундаменту построения ГИС.

Четвертая часть "Организация данных в геоинформационных системах" посвящена рассмотрению проблем системной организации данных. Особое внимание уделено двум распространенным моделям организации данных - георегиональной модели и объектно-ориентированной модели данных.

Пятая часть "Основы геопространственного анализа" является заключительным модулем, который знакомит читателя с тем, что называют "сердцем" ГИС. Эта часть включает общую характеристику и методологию геопространственного анализа. На основе распространенной классификации рассмотрены базовые аналитические средства, которые являются опорой геопространственного анализа - функции измерений, выбора данных, классификации, оверлейные функции, функции окрестности, функции связности.

Для создания пособия использован обширный набор литературных источников, многие из которых сами по себе могут быть интересными объектами изучения.

Предлагаемое пособие написано на основе многолетнего опыта преподавания курса "Основы ГИС" для студентов специальности

"Геоинформационные системы и технологии", а также для будущих градостроителей, экологов, менеджеров в Харьковской национальной академии городского хозяйства. Использован опыт академии разработок и внедрения ГИС-проектов для города Харькова и других городов. Пособие может быть полезным студентам специальностей компьютерных наук, а также специалистам в областях управления территориями, земельными ресурсами и недвижимостью, коммунальным хозяйством, транспортной инфраструктурой, в областях энергетики, экологии, в областях разработки, создания или исследования сложных пространственных комплексов.

Часть 1

**ВВЕДЕНИЕ
В ГЕОГРАФИЧЕСКИЕ
ИНФОРМАЦИОННЫЕ СИСТЕМЫ**

Раздел 1.1

КОНЦЕПЦИЯ СИСТЕМ

Определение термина "Географическая информационная система" требует в первую очередь рассмотрения понятия "Система" и его основных концептов. Это важно потому, что понятийная база теории систем, общие закономерности развития и функционирования систем являются фундаментом знаний о географических информационных системах.

1.1.1 Сложность и взаимозависимость частей нашего мира

Прогресс в изучении окружающего нас мира приводит к представлениям о все большей *сложности* и *взаимозависимости* его частей.

Сложность – это совокупность огромного числа различных объектов. Сложность определяется обилием связей. Следует различать простую и трудную проблемы.

- Простая проблема имеет одно решение (например, развязывание узла на веревке может быть трудным).
- Сложная проблема имеет много возможных решений [1].

Проблемы *не существуют изолированно*, являются ли они социальными, политическими или экономическими. Пример: экономика зависит от энергетики и др. ресурсов, энергоресурсы – от политики, политика – от силы; и наоборот, политика зависит от экономики. Проблемы не могут быть выделены из целого, объяснены по отдельности, а затем интерпретированы для объяснения целого.

Среда, в которой возникают проблемы, сама по себе не является статичной. Она *динамична*, так как всегда изменяется, подвергаясь внешним и внутренним воздействиям. Среда меняется, взаимодействует вместе со своими проблемами и их решениями в физическом и концептуальном пространстве. Среда меняется во времени, так как на нее влияют события.

В последнее время для многих областей науки и техники *характерен переход от разработки отдельных объектов к созданию сложных комплексов*. Вырос масштаб проблем, комплексность и сложность проблем, стоимость их решения. Потребность решения сложных проблем породила необходимость в *универсальной*

методологии, которая позволяет с единых позиций выполнить анализ и синтез сложных объектов. В связи с этим получила развитие методология, в основу которой положена *концепция систем*.

1.1.2 Исторический аспект

Первые представления о системах возникли в античной философии [1]. В древнегреческой философии и науке (Эвклид, Платон, Аристотель) разрабатывалась идея системности знания. Философы всех эпох по-своему использовали понятия системы. Принцип системной природы знаний разрабатывались в классической философии (Спиноза, Шеллинг, Гегель...).

В 19 веке понятия системы уже используется во многих областях конкретно-научного знания и общенаучного знания. Ф. Энгельс в работе "Диалектика природы" пишет: "*Вся доступная нам природа образует некоторую систему, некоторую совокупную связь тел, причем мы понимаем здесь под словом "тело" все материальные реальности, начиная от звезды и кончая атомом.... В том обстоятельстве, что тела находятся во взаимной связи, уже заключено то, что они взаимодействуют друг на друга, и это их взаимное воздействие друг на друга и есть именно движение. Движение несотворимо и неуничтожимо. Этот вывод стал неизбежным, лишь только люди познали вселенную, как систему, как взаимную связь тел*".

В 20 веке многие крупные ученые внесли свой вклад в развитие системных представлений, в их числе В.И. Вернадский, А.А. Богданов и другие.

Объединение системных представлений в особую область знания связывают с именем австрийского биолога *Людвига фон Бергаланфи*, который в конце 40-х годов 20 века выступил с программой построения "*общей теории систем*". Дополнительно к этой программе в 50-70 годах был выдвинут ряд общесистемных концепций и определений понятия система. Неоценимый вклад в системные исследования внесли американские ученые М. Месарович, Р. Акоф, английский кибернетик У. Р. Эшби, советские ученые А. И. Уемов, И. В. Блауберг, В. Н. Садовский, Э. Г. Юдин, А. Н. Аверьянов и другие.

1.1.3 Определения понятия "Система"

Система – универсальное понятие, пригодное для обозначения разнообразных объектов любой природы. Известен ряд определений системы. Вот некоторые из них:

- Система – есть комплекс взаимодействующих элементов [2]
- Система — (от греческого *systema* - целое, составленное из частей; соединение) – это совокупность элементов, находящихся в отношениях и связях друг с другом, которые образуют определенную целостность, единство [3, с.610].
- Система – есть ограниченное множество взаимодействующих элементов. [4, с.43]
- Система – есть множество взаимосвязанных элементов [5]
- Система – есть то, что решает проблему. [6]

Понятие "система" имеет теснейшую взаимосвязь с понятиями целостности, структуры, связи, элемента, отношения, подсистемы. Наличие связей и отношений между элементами системы является определяющим, ключевым.

Практически каждый объект, процесс, явление может быть рассмотрен как система. Примеры систем:

- Солнечная система, транспортная система, система счисления, система Станиславского, экономическая система, геологическая система, система понятий, нервная система, система химических элементов Менделеева, система математических уравнений...
- Системами являются язык, комбайн, мозг, телефонная сеть, вычислительная машина, человек, академическая группа, факультет, институт, система образования, общество людей,...
- Город – сложная динамическая социальная система. Эта система может быть представлена взаимодействующими подсистемами городского хозяйства. Специалист должен знать и уметь анализировать структуры, взаимосвязи, критерии оценки функционирования систем городского хозяйства для принятия решений при планировании, проектировании, создания и управления систем.
- Программный продукт.
- Географическая информационная система.

Понятие система имеет чрезвычайно широкое применение. Рассмотрение объекта или совокупности объектов, среды как системы позволяет на основании общих закономерностей развития и функционирования систем конструировать сложные системы,

определять взаимосвязи между элементами системы, эффективно решать сложные проблемы.

1.1.4 Системные исследования

Под системными исследованиями понимается вся совокупность научных и технических проблем, которые сходны в понимании и рассмотрении исследуемых ими объектов как систем, т.е. множеств взаимосвязанных элементов, выступающих как единое целое. Системные исследования представлены тремя системными областями:

- *Системный подход (СП)* – общенаучное методологическое направление, разрабатывающие методы и способы исследования сложноорганизованных объектов; СП основанный на рассмотрении объектов изучения в виде систем. СП – сердцевина системных исследований.
- *Системный анализ (СА)* - методология, ориентированная на решения крупных проблем практически прикладного характера, основанная на концепции систем. В центре СА, операция сравнения альтернатив решения проблем и выбора альтернативы.
- *Общая теория систем (ОТС)* – научная дисциплина, которая изучает различные явления, отвлекаясь от их конкретной природы, и считается лишь на *формальных взаимосвязях* между различными составляющими их факторами, между наблюдениями, признаками и свойствами. Задачи ОТС:
 - разработка обобщенных моделей систем;
 - построение логико-методологического аппарата;
 - создание обобщенных теорий систем различного типа.

1.1.5 Элементы систем

Система или любая ее часть может быть расчленена на элементы.

Под элементом системы понимают такую ее часть, которая не подлежит дальнейшему расчленению, является надежной с точки зрения процесса функционирования системы. Элемент - это составная неделимая часть целого.

Элементы системы обладают определенными свойствами.

Элементами могут быть предметы, явления, знания, методы.

Примеры элементов систем:

- химические элементы системы Менделеева;

- планеты солнечной системы;
- студенты академической группы;
- транспортные средства, дороги, люди, грузы транспортной системы.

При рассмотрении системы "*Предприятие*" ее элементами могут быть отделы заводоуправления, цехи, склады и т.д. При рассмотрении системы "*Цех*" элементами будут станки, бригады рабочих, контора цеха и т.д. При рассмотрении системы "*Человек*" элементами будут органы дыхания, кровообращение, пищеварение, движение и т.д.

Следовательно, процесс расчленения системы на элементы и само понятие элемента является *относительным и условным*.

1.1.6 Системные связи

Понятие связи – одно из важнейших, фундаментальных научных понятий. Системы связаны между собой. Элементы системы связаны между собой.

Связь – это взаимообусловленность существования объектов, явлений, процессов, распределенных в пространстве и времени. Связь – есть то, что соединяет элементы, объекты и свойства в системном процессе. Связи определяют следования процессов.

Система есть множество взаимосвязанных элементов. Существует два критерия системности:

- каждый элемент системы связан, по крайней мере, с одним другим элементом системы;
- каждое подмножество элементов системы связано с каждым другим подмножеством системы.

Графически связь отношения одного элемента к другому изображается направленной линией (\rightarrow).

Различают [5] между элементами a, b, c, d следующие типы связей:

- по направлению действия:
 - Прямые ($a \rightarrow b, b \rightarrow c$);
 - Обратные ($c \rightarrow b, b \rightarrow a$);
- по контакту:
 - непосредственные ($a \rightarrow b, b \rightarrow c$);
 - опосредованные ($a \rightarrow c, b \rightarrow c$);
- по следованию:
 - последовательные ($a \rightarrow b \rightarrow c$);
 - параллельные ($b \rightarrow d, b \rightarrow c$);

- по отношению к системе:
 - внешние;
 - внутренние;
- по силе:
 - жесткие;
 - корпускулярные;
- по субстрату, содержанию связи:
 - вещественные;
 - энергетические;
 - информационные.

1.1.7 Понятие "Структура"

Важнейшим системным понятием является понятие "структура". Термин "структура" происходит от латинского *structure* – строение, расположение, порядок.

Структура – совокупность устойчивых связей системы, обеспечивающих его целостность и тождественность самому себе, т.е. сохранение основных свойств при различных внешних и внутренних изменениях [3]. Структура – это устойчивая форма связей между элементами системы.

Структура выражает лишь то, что остается устойчивым, относительно неизменным при различных преобразованиях системы.

Выделение структуры объекта – является важнейшей составляющей изучения объекта как системы. Основные *шаги выделения* структуры:

- 1) выделение множества элементов системы;
- 2) определение пространственных расположений элементов пространства;
- 3) выделение связей (отношения) между элементами, которые являются инвариантными при определенных изменениях системы.

Любая система имеет иерархическую структуру.

1.1.8 Системный эффект

Системный эффект выражен постулатом Аристотеля

"Целое – больше суммы его частей"

Это означает, что "Свойства предметов и способы действий на высших уровнях не могут быть выражены при помощи суммирования свойств и действий их компонентов, взятых изолированно. Если, однако, известен ансамбль компонентов и существующие между ними отношения, то высшие уровни могут быть выведены из компонентов" [2]. Пример 1: в учебном процессе "сумма знаний не есть система знаний"; но если отдельные знания системно организованы и получены, они образуют систему знаний. Пример 2: прибыль после слияния двух компаний превосходит сумму прибылей этих компаний до объединения.

По отношению к действиям системный эффект обозначают термином Синергия. (от греч. Synergos — вместе действующий). Синергия — это комбинированное воздействие двух или более факторов, характеризующееся тем, что их объединённое действие существенно превосходит эффект каждого отдельно взятого компонента и их суммы.

Благодаря системному эффекту наш мир гораздо больше, чем простая совокупность изолированных и автономных элементов. Для познания в равной степени необходимы не только элементы, но и связи между ними.

1.1.9 Системные принципы

При рассмотрении объекта как системы должны соблюдаться следующие основные системные принципы [3]:

- 1) принцип целостности -
принципиальная несводимость свойств системы к сумме свойств составляющих ее элементов и невыводимость из последних свойств целого;
- 2) принцип структурности -
возможность описания системы через установленные ее структуры, т.е. сети связей и отношений системы; обусловленность поведения системы не столько поведением ее отдельных элементов, сколько свойствами ее структуры;
- 3) принцип взаимосвязанности системы и среды -

система формирует и проявляет свои свойства в процессе взаимодействия со средой, являясь при этом ведущим активным компонентом взаимодействия;

- 4) принцип иерархичности - каждый компонент системы в свою очередь может рассматриваться как система или подсистема, а исследуемая в данном случае система представляет собой один из компонентов более широкой системы; сложные системы содержат много разных иерархий;
- 5) принцип множественности описания каждой системы - в силу принципиальной сложности системы ее адекватное познание требует построения множества различных моделей, каждая из которых описывает лишь определенный аспект системы.

1.1.10 Формы и способы описания систем

Возможны [7, с. 121] следующие *формы* описания систем:

- математическая форма,
- графическая форма,
- текстовая форма.

Используют два *способа* описания систем:

- внутреннее описание систем,
- внешнее описание систем.

Внутреннее описание представляет систему через множество элементов и связей.

- Графическая форма – в виде графов, связанных таблиц;
- Математическая форма – через множество параметров Q_i , называемых переменными состояния, и математическими зависимостями. В общем случае поведение системы, зависимое от времени t , описывается системой n дифференциальных уравнений первого порядка:

$$\frac{dQ_n}{dt} = f_i(Q_1, Q_2, \dots, Q_n) \quad (1.1.1)$$

При внешнем описании система принимается за «черный ящик». Черным ящиком называется система, внутреннее содержание которой неизвестно. Существенными являются внешние связи системы, которые обеспечивают взаимодействие между системой в целом и внешней средой (Рис.1.1.1). Внешние связи можно рассматривать

также как связи между элементами данной системы с элементами других систем, составляющих внешнюю среду.

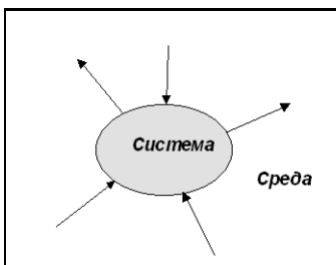


Рис. 1.1.1 – Отношения системы и среды

Отношения со средой и другими системами графически изображаются в виде блок-схем и диаграмм (Рис. 1.1.2).

Описание системы производится в терминах Входов и Выходов:

Вход > Черный ящик > Выход

При внешнем описании систем используются системные объекты:

- *Вход* – то, что изменяется при протекании данного процесса;
- *Выход* – результат процесса;
- *Процесс* – то, что переводит вход в выход;
- *Связи* – есть то, что соединяет системные объекты в системном процессе;
- *Ограничения* – состоят из целей и принуждающих связей;
- *Обратная связь* – определяется как системный процесс, который сравнивает критерии и выход.

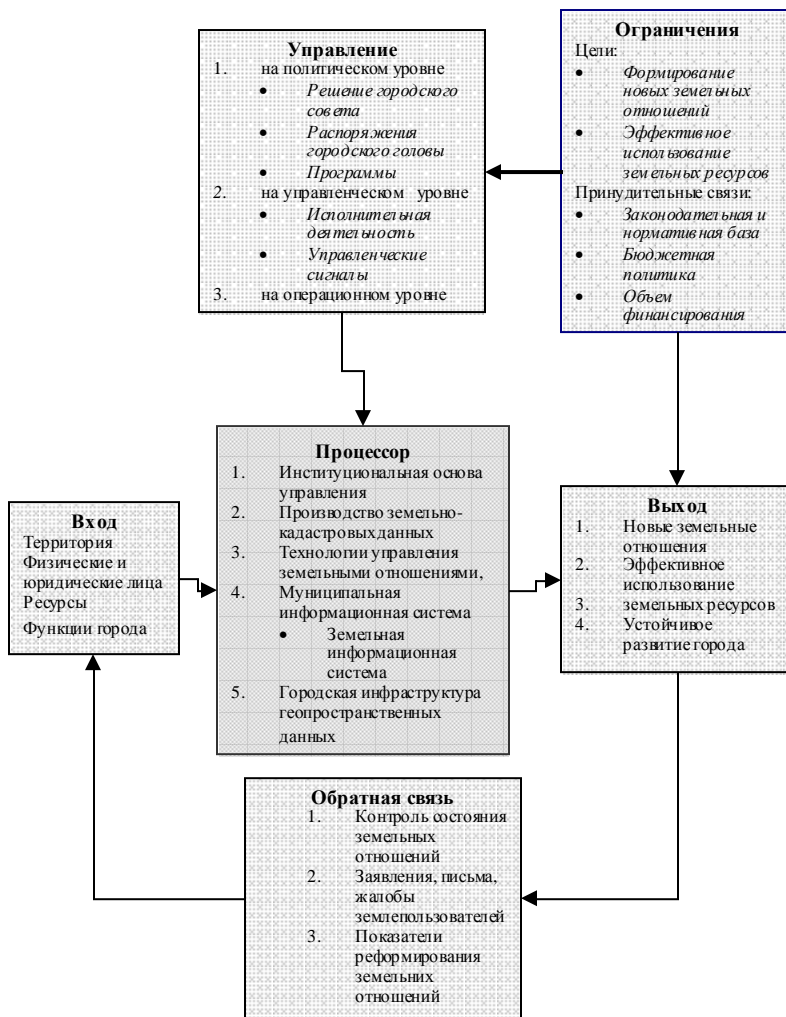


Рис. 1.1.2 – Пример графического описания системы управления земельными ресурсами города [8].

1.1.11 Классификация систем

Существенным аспектом раскрытия содержания понятия система является выделение различных типов систем. Системы подразделяются:

- 1) по свойствам:
 - материальные (вещественные);
 - неорганической природы (физические, геологические);
 - живые (биологические системы, организмы, популяции, виды, экосистемы);
 - социальные (социально-экономические структуры обществ);
 - абстрактные (понятия, гипотезы, теории, научные знания, формализованные);
- 2) по происхождению:
 - естественные;
 - искусственные;
- 3) по характеру перехода к другому состоянию:
 - статичные;
 - динамичные;
- 4) по знанию значений переменных:
 - детерминированные;
 - вероятностные (стохастические);
- 5) по характеру взаимоотношений системы и среды:
 - закрытые;
 - открытые;
- 6) по сложности:
 - простые;
 - сложные;
 - очень сложные.

Примеры классификации систем приведены в табл. 1.1.1.

В руководстве Гради Буча [9, с.13] представлены пять признаков сложной системы.

- 1) Сложные системы часто являются иерархическими и состоят из взаимозависимых подсистем, которые в свою очередь также могут быть разделены на подсистемы, и т.д., вплоть до самого низкого уровня.
- 2) Выбор, какие компоненты в данной системе считаются элементарными, относительно произволен и в большой степени оставляется на усмотрение исследователя.

- 3) Внутриконтентная связь обычно сильнее, чем связь между контентами. Это обстоятельство позволяет отделять "высокочастотные" взаимодействия внутри контентов от "низкочастотной" динамики взаимодействия между контентами.
- 4) Иерархические системы обычно состоят из немногих типов подсистем, по-разному скомбинированных и организованных.
- 5) Любая работающая сложная система является результатом развития работавшей более простой системы... Сложная система, спроектированная "с нуля", никогда не заработает. Следует начинать с работающей простой системы.

Таблица 1.1.1 Примеры классификации систем.

Системы	Простые	Сложные	Очень сложные
Детерминированные (определенные)	Оконная задвижка Проект механической мастерской	ЭВМ Автоматизация Двигатель	
Вероятностные (стохастические)	Подбрасывание монет Движение медузы Контроль качества продукции	Хранение запасов Условный рефлекс Прибыль промышленного предприятия	Экономика Мозг Город

1.1.12 Контрольные вопросы и задания для самостоятельной работы

- 1) Дайте определение понятия Система.
- 2) Дайте определение понятий Элементы систем, Системные связи, Структура системы.
- 3) Какие основные принципы должны соблюдаться при рассмотрении объекта как система?
- 4) Какие существуют формы и способы описания систем?
- 5) Какие бывают типы систем?

Раздел 12

ОБЩАЯ ХАРАКТЕРИСТИКА ГЕОГРАФИЧЕСКИХ ИНФОРМАЦИОННЫХ СИСТЕМ

1.2.1 Область ГИС

В последние два десятилетия получил широкое распространение особый класс информационных систем, который называют *географические информационные системы*, или сокращенно *геоинформационные системы*, или аббревиатурой "*ГИС*" (Geographical Information System – GIS).

Область ГИС очень быстро развивается и захватывает все новые сферы жизни и деятельности. Геоинформационные системы стали играть существенную роль в деятельности человечества. Причины такого успеха ГИС объясняется следующим:

- во-первых, большой эффективностью решений сложных проблем средствами ГИС;
- во-вторых, огромным множеством областей применения ГИС потому, что ГИС работает с пространственными данными, которые являются частью нашей повседневной жизни;
- в-третьих, стали доступными для массового пользователя как мощные персональные компьютеры, на которых возможна реализация ГИС, так и сложное программное обеспечение ГИС, снабженное легким пользовательским интерфейсом.

В последние 15-20 лет геоинформационные системы достигли в развитии уровня коммерциализации. В настоящее время сформировалась ГИС-индустрия, оборот которой измеряется миллиардами долларов. Ежегодный объем продаж только программного обеспечения ГИС в 2001 году составлял 1,1 млрд. долларов, в 2009 году превышает 2,8 млрд. долларов (увеличивается в среднем на 12% в год).

В ГИС-движение вовлечены миллионы пользователей в более чем в 120 странах мира, которые образуют ГИС-сообщество. Специалисты по работе с ГИС в странах Запада пользуются повышенным спросом.

О высоком интересе к этой технологии свидетельствует большое число ежегодно проводимых конференций и симпозиумов.

Наблюдается рост числа участников конференций. Особенно поразительный рост на ежегодной итоговой конференции для пользователей программного обеспечения ESRI (института исследования систем окружающей среды - мирового лидера в области разработки программного обеспечения ГИС). В 1998 году в этой итоговой конференции участвовало 8 тыс. человек из 90 стран мира; в 1999 году – 9; в 2001 г. – 11; в 2002 г. – 12; в 2006 г. - 13, в 2008 г. – более 14 тыс. пользователей из 124 стран мира. ГИС изучают не только в университетах, в системах повышения квалификации, но уже и в школах, колледжах.

Геоинформационные системы влияют на многие аспекты нашей жизни. ГИС интегрируются в образование, бизнес, производство и деятельность правительств.

Управление сложными территориальными системами, такими как современный город, требует не только высокой квалификации администрации, но и хорошего информационного обеспечения процесса управления. Выработка обоснованных и эффективных решений возможна только на основе соответствующей обработки и анализа огромных объемов информации о сложных системах и процессах города. Это требует привлечения эффективных средств, базирующихся на компьютерных технологиях. Внедрение ГИС в управление городом означает переход на новый, более эффективный уровень управления.

ГИС в науке помогает интегрировать научные области, ломая барьеры между биологией и геофизикой, между экономикой и демографией. Этому способствует моделирование процессов в реальном времени, мощный инструментальный пространственного анализа, высокое пространственное разрешение данных.

Растущий интерес к ГИС иллюстрирует рост потребности в пространственных данных высокого разрешения, получаемых из космоса. Для удовлетворения этих растущих потребностей запущены новые спутники: в 2001 г. спутник Quick-Bird с разрешением 0,61 м. панхроматических данных, в 2007 г. - World View-1 с разрешением 0,50 м., в 2008 г. - GeoEye-1 с разрешением 0,41 м.

В прошедшие годы усилия концентрировались на создании пространственных баз данных, на накоплении данных. Это в свою очередь привело к необходимости разработки проблемы распространения геопро пространственных данных и обмена данными, в том числе создания стандартов на инфраструктуру геопро пространственных данных. В Правительственном распоряжении 1994 г. о национальной инфраструктуре геопро пространственных данных

Президента США Уильяма Дж. Клинтона геоинформационным технологиям придается следующее значение: "Географическая информация является важнейшим условием обеспечения экономического развития, совершенствования управления природными ресурсами и охраны окружающей среды" [10].

Дальнейшее развитие ГИС обуславливают высокая динамика жизни, рост населения, урбанизация, уменьшение и ограничение природных ресурсов, глобализация. Двадцать первое столетие – период широкого внедрения средств для интеграции информации. В информационном обществе ГИС становится одной из важнейших технологий интеграции и коммуникации информации.

1.2.2 Определение ГИС

Когда спросили известного специалиста в области геоинформатики Нэнси Тоста о перспективе ГИС в новом тысячелетии, она ответила: "Я думаю, что в ближайшие 10 лет прекратят задавать вопрос: "Что такое ГИС?". В этом ответе заложен глубокий смысл в контексте развития этой области.

Одни специалисты используют проблемно-ориентированные подходы и передовые новые методологии, используя ГИС как средство в более чем 70 дисциплинах. Другие рассматривают ГИС как дисциплину или как тему в информационных технологиях или компьютерных науках. Поэтому еще можно встретить определения типа "ГИС – это программное обеспечение", или "ГИС – это электронная карта", или "ГИС – это база данных", которые существенно отличаются от распространенного понимания этой области человеческой деятельности.

При любом взгляде в рассмотрение должны быть включены определенные фундаментальные концепции, типа системы, картографические проекции, масштабы, структуры пространственных данных, ввод, хранение, преобразование и точность данных, метаданные, пространственный анализ, проектирование, управление и внедрение ГИС.

Известными специалистами в области ГИС опубликовано несколько десятков определений ГИС, которые в совокупности определяют устоявшиеся характеристики ГИС. Вот некоторые из них.

Michael H. DeMers: "Геоинформационные системы - это инструменты для обработки пространственной информации, обычно

явно привязанной к некоторой части земной поверхности и используемые для управления ею. Как и в случае с географией, термин Географическая информационная система (ГИС) трудноопределим. Он представляет интеграцию многих предметных областей. Поэтому нет общепринятого определения ГИС" [11, с.8].

Michael F. Goodchild: "Существует два отдельных мнения относительно вопроса "Что есть ГИС?":

- 1) ГИС – это реальное приложение, включающее аппаратное обеспечение, данные, программное обеспечение и людей, необходимое для решения проблем (ГИС-приложение);
- 2) ГИС – это тип программного обеспечения, продаваемого разработчиками программного обеспечения.

Мы будем фокусировать на первом определении" [12].

NCGIA: "Широко распространенное определение ГИС: ГИС – это система аппаратных средств, программного обеспечения и процедур, способствующая управлению, манипулированию, анализу, моделированию, представлению и отображению географических данных для решения комплексных проблем планирования и управления ресурсами" [13].

Эти и ряд других известных определений позволяют синтезировать определение, которое соответствует наиболее распространенному пониманию этой области:

"ГИС - это система, которая:

- *во-первых, представляет собой комплекс взаимодействующих пяти компонентов, состоящий из компьютерных средств, программного обеспечения, географических данных, регламента и пользователей;*
- *во-вторых, выполняет функции ввода, интегрирования, хранения, обработки, анализа, моделирования и визуализации географической информации".*

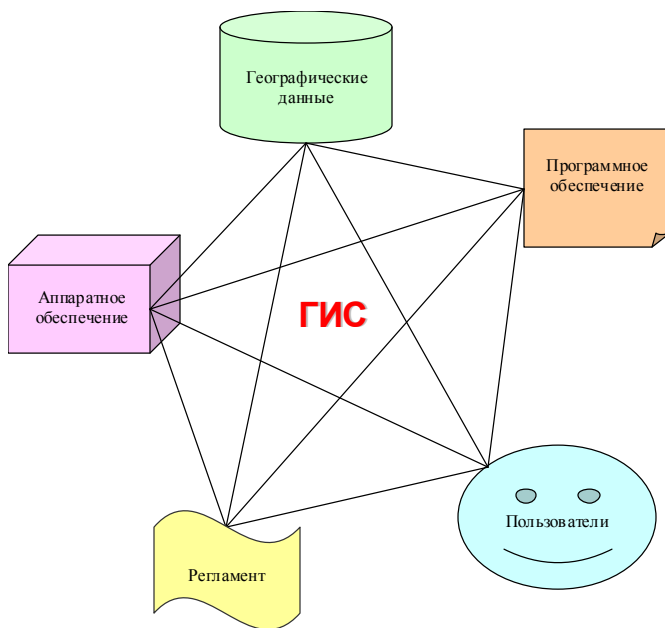


Рис. 2.1.1 Графическое представление геоинформационной системы

Известны также некоторые оригинальные определения:

- ГИС - это возможность нового взгляда на окружающий нас мир;
- ГИС - это инструмент видения мира во взаимосвязях;
- ГИС – это мощное средство интеграции и организации информации.

1.2.3 Общая характеристика компонентов ГИС

1.2.3.1 Аппаратные средства ГИС

Эффективная работа современных ГИС возможна только на основе мощной технической компьютерной поддержки.

Аппаратные средства ГИС - это отдельные компьютеры или системы компьютеров с сетевым оборудованием плюс компьютерная

периферия, к которой относят принтеры, плоттеры, сканеры. Аппаратные средства ГИС должны обеспечивать возможность:

- ввода данных из разных источников,
- хранения, манипулирования, обработки больших массивов данных,
- выполнения сложных операций пространственного анализа больших массивов данных,
- качественного представления данных на монитор или на бумажные носители,
- эффективное взаимодействие компонентов.

Лавинообразному внедрению ГИС способствовали достижения в области производства персональных компьютеров на основе процессоров класса Pentium, доступных для массового использования. Аппаратные средства развиваются очень быстро и требуют специального освещения вне этого учебного пособия.

Аппаратные средства ГИС совместно с техническими средствами позиционирования (электронными геодезическими станциями, приемниками спутниковых систем позиционирования), средствами дистанционного зондирования Земли интегрируются в технические комплексы, позволяющие эффективно получать и обрабатывать геопространственную информацию.

1.2.3.2 Программное обеспечение ГИС

Программное обеспечение ГИС можно представить как совокупность взаимосвязанных концентрических оболочек. Ядро такой совокупности составляет системное программное обеспечение, в том числе операционная система, система управления базой данных и др.

Оболочкой ядра является базовое инструментальное программное средство ГИС, которое обеспечивает выполнение всех функций: ввод, интегрирование, хранение, обработка, анализ и презентация географической информации. Базовое инструментальное программное средство ГИС может быть представлено совокупностью программных продуктов, реализующих отдельные функции. Но если программный продукт не выполняет все указанные функции ГИС, он рассматривается как специализированное ГИС-приложение, предназначенное для удовлетворения специфических запросов пользователей.

К настоящему времени в мире разработано сотни коммерческих программных продуктов ГИС разного качества и функциональных возможностей. Разобраться в этом море не просто даже специалистам. Ошибка в выборе базового программного продукта связана с большими потерями финансовых и временных ресурсов.

При всем разнообразии конкретных систем в них можно выделить инвариантное технологическое ядро, независимое от предметной области, ни от проблемной ориентации, ни от приложений вообще, что позволяет говорить о некоторой геоинформационной технологии как технологии сбора, ввода, интегрирования, хранения, обработки, анализа и презентации географической информации, которая выкристаллизовалась из опыта реализации и эксплуатации различных геоинформационных систем.

По данным маркетингового анализа индустрии ГИС, традиционно проводимого разными фирмами, в том числе фирмой Daratech, Inc., определяются ежегодно доли рынка наиболее распространенных программных продуктов ГИС. В 2009 году эти доли составили для ESRI 30%, Intergraph 18%, Smallworld (GE) 8%, ... (Рис.1.2.2).

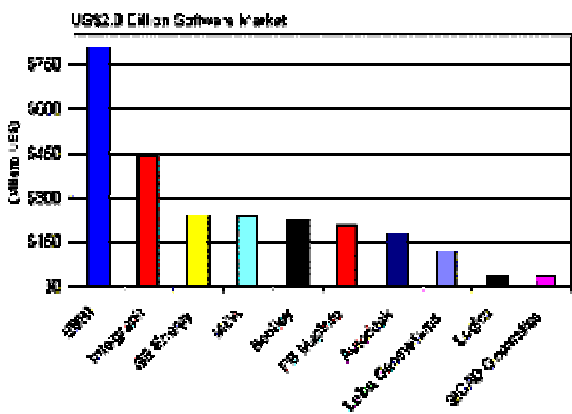


Рис. 1.2.2 - Мировые объемы продаж программного обеспечения ГИС в 2009 году по данным Daratech, Inc.

Эти исследования подтверждают, что на протяжении многих лет мировым лидером в разработке и распространении программного обеспечения ГИС является *Институт исследования систем окружающей среды* (Environmental Systems Research Institute - ESRI). В среднем третья часть продаж программного обеспечения ГИС приходится на долю ESRI и составляет 37-29 %, что говорит о многих

преимуществах и привлекательных свойствах программного обеспечения этой фирмы. Это является основанием того, что обучение студентов целесообразно проводить на основе наиболее передовых технологий ESRI.

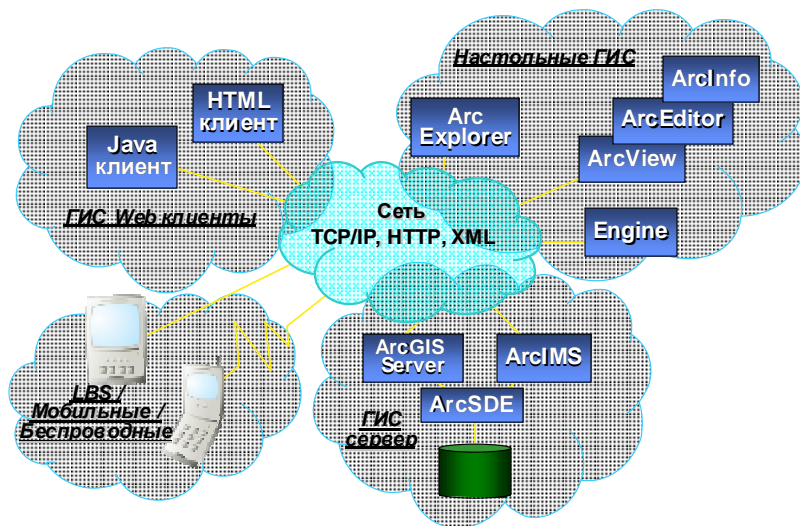


Рис. 1.2.3 – Компоненты системы программных продуктов ArcGIS 9.x

Компанией ESRI создана мощная система программных продуктов под названием ArcGIS 9.x (Рис. 1.2.3), которая является оптимальным решением для построения корпоративной ГИС, фундамента информационной системы эффективного управления крупными государственными и коммерческими организациями.

Фундаментальная архитектура ArcGIS 9.x обеспечивает ее использование во многих прикладных сферах и на разных уровнях организации работы: на персональных компьютерах, на серверах, через Web или в полевых условиях. ArcGIS 9.3 включает:

- *настольные ГИС* (Arc View, ArcEditor, ArcInfo, Desktop Extensions, ArcGIS Engine, ArcExplorer, ArcReader), которые позволяют эффективно выполнять и моделировать операционный процесс, выполнять продвинутый пространственный анализ, визуализировать результаты на картах высокого качества;

- *серверные ГИС* (ArcGIS Server, ArcGIS Server Extensions, ArcGIS for AutoCAD) , которые позволяют распределять карты, модели и инструменты другим пользователям в организации так, что они хорошо соответствуют их процессам;
- *мобильные ГИС* (ArcGIS Mobile, ArcPad), которые предоставляют возможность исполнителям в полевых условиях получать, хранить, обновлять, манипулировать, анализировать и отображать географическую информацию;
- *интернет ГИС* (ArcGIS Online), которые обеспечивают общую платформу для пользователей ArcGIS, чтобы находить, распределять и создавать ГИС приложения, сервисы и информационные ресурсы;
- *инструменты разработчика* (Developer Tools), которые обеспечивают интерфейсом программирования приложений, программами и ресурсами так, что разработчик может создать инновационное ГИС-решение для настольной, мобильной, серверной или интернет ГИС;
- *специализированные ГИС*:

ESRI Business Analyst - представляет собой доступное и масштабируемое семейство продуктов, которые позволяют малым предприятиям и предприятиям получить доступ к большому объему данных и увеличить свою прибыль на инвестиции в этом экономическом климате;

ArcLogistics - представляет собой комплексное решение для создания оптимизированных маршрутов и решения задач составления расписаний.

ArcLogistics Navigator - представляет собой GPS-управляемый транспортный навигационный продукт, который позволяет мобильным профессионалам пребывать на маршруте и в установленные сроки.

MapIt - это программное обеспечение и онлайн-услуги, которые позволят вам создать простые карты из ваших корпоративных данных;

PLTS ArcGIS используют для всех типов организаций в целях улучшения своих рабочих процессов и для удовлетворения своих самых сложных карт и диаграмм потребностей.

Situational Awareness for ArcGIS представляет собой решение, которое обеспечивает геопространственную основу для немедленных и долгосрочных потребностей ситуационной осведомленности.

Следует отметить, что в настоящее время также имеет большое распространение ранее созданные фирмой ESRI программные продукты ArcView GIS версии 3.x, относительно простые в использовании и эффективные для многих задач. Важно то, что данные, созданные в ArcView GIS 3.x, воспринимаются в ArcGIS 9.x.

По сравнению с ESRI, меньшее распространение программных продуктов ГИС имеют другие фирмы. Среди них по убыванию доли в мировом рынке следует назвать фирмы и соответствующие программные продукты ГИС:

- | | |
|---|--|
| ▪ Intergraph Corp. | MGE
(Modular GIS
Environment),
GeoMedia
Smallworld GIS |
| ▪ GE Network Solution
(General Electric Company) | |
| ▪ Bentley Systems Inc. | MicroStation |
| ▪ Mapping Information Systems Corp. | MapInfo Professional |
| ▪ Autodesk Inc. | AutoCAD,
Autodesk Map,
Autodesk Survey |

В России и Украине также имеют распространение отечественные программные продукты ГИС:

- | | |
|--|---------------------|
| ▪ Институт географии
Российской академии наук | GeoDraw / Geo Graph |
| ▪ КБ "Панорама", МО, Россия | Panorama,
Карта |
| ▪ ГП "Геосистема", Винница, Украина | Digitalis,
Delta |

1.2.3.3 Географические данные

Геоинформационные системы отличаются от других информационных систем тем, что они имеют богатые возможности работы с геопространственными данными.

Географические данные (синоним - геопространственные данные) описывают любые объекты, имеющие локализацию в реальном земном пространстве. Они состоят из двух взаимосвязанных частей: пространственные данные, описывающие местоположение, форму и

размеры объекта, и семантические данные, описывающие содержательные характеристики объекта.

Местоположение в реальном земном пространстве описывается геодезическими координатами. Формы и размеры объектов определяются посредством координат. Пространственные модели объектов создаются на основе картографических образов, с которыми мы все знакомы. Поэтому картографическое представление реального мира играет существенную роль в геоинформационных технологиях. Внешне вы видите на дисплее компьютера цифровую карту. Сложные функции пространственного анализа реализуются посредством обработки таких карт.

Чрезвычайно важным обстоятельством является то, что объем пространственных данных, необходимых для управления территорией, составляет 70 – 90 % по оценкам научных исследований, выполненных разными организациями. Это означает, что ГИС обрабатывает большую часть информации, используемой как при управлении территорией, так и ее компонентами. ГИС поэтому играет ведущую роль в управлении территориями. Класс геоинформационных систем специально создан для решения актуальных проблем земного пространства, возникающих в самых разных сферах человеческой деятельности.

Создание географических данных - наиболее трудоемкая, дорогостоящая и ответственная часть ГИС. По мировым оценкам стоимость географических данных составляет до 80-90 % стоимости системы. Поэтому при создании ГИС исключительно важными являются задачи: 1) сохранности данных длительный период времени - до 80 лет, 2) возможности обмена данными между системами, существующими сегодня, 3) передачи данных в будущие системы в условиях, когда программно-техническое обеспечение существенно меняется через 2-3 года. Определяющими здесь являются форматы данных. Они должны быть или общепринятыми или такими, что позволяют делать обмен данными.

Географические явления и феномены бесконечно сложны и разнообразны. Мы можем рассматривать их с любой степенью детальности. Поэтому для абсолютно точного описания сущностей реального мира потребовались бы бесконечно большие базы данных. Для того, чтобы быть обработанными средствами вычислительной техники, данные должны быть редуцированы до конечных размеров и организованы особым образом.

1.2.3.4 Регламент

Регламент ГИС устанавливает жесткие правила системной организации данных, информационной деятельности и технологии работ. Разработка регламента функционирования системы ведется в период планирования ГИС.

Регламент ГИС требует использования ряда стандартов. Стандарты устанавливают, как данные определяются, хранятся и перемещаются между системами и приложениями. Эти стандарты могут быть сконструированы так, чтобы обеспечить оптимальный баланс между совместным и индивидуальным использованием посредством определения минимальных требований для обмена данными.

1.2.3.5 Пользователи

Работа ГИС невозможна без разработчиков, обслуживающего персонала и пользователей. От них зависит эффективность ГИС. Обязательным условием реализации ГИС является наличие важнейшего компонента – *думающих пользователей*, которые имеют подготовку для изучения земного пространства, работы с географическими данными, умеют работать в среде ГИС-приложений, создавать их и поддерживать, имеют определенные знания из предметной области приложения ГИС. Для создания ГИС нужны разработчики, а для функционирования – пользователи, причем количество последних составляет подавляющее большинство.

1.2.4 Общая характеристика отличительных функций ГИС

Ранее было определено, что ГИС выполняет функции ввода, интегрирования, хранения, обработки, анализа, моделирования и визуализации географической информации. Специфическими функциями, посредством которых ГИС выделяется в особый класс информационных систем, являются интегрирование, анализ и визуализация географической информации.

1.2.4.1 Интегрирование

В идеологии ГИС заложена идея интегрирования.

Геоинформационные системы отличаются от других информационных систем именно тем, что обладают эффективными возможностями интегрирования разноплановой пространственной информации, связанной с реальным земным пространством. Это означает, что любые данные могут быть интегрированы в одну систему, если эти данные имеют или могут иметь пространственную привязку в реальном земном пространстве.

Например, одна муниципальная ГИС интегрирует информацию о градостроительстве, населении, землепользовании, жилым и нежилым фондом, городских инженерных системах водоснабжения, водоотведения, теплоснабжения, газоснабжения, электроснабжения, транспортной инфраструктуры, об экологической обстановке, правонарушениях... Это дает возможность управлять, например, подсистемами города комплексно и во взаимосвязях, определять влияние одних факторов на другие.

В общем случае интегрированные географические данные предоставляют возможность:

- реализовать комплексный подход при анализе совокупности сущностей реального мира,
- определять влияние, зависимости, связи одних объектов и явлений реального мира на другие,
- вырабатывать предложения для принятия объективных управленческих решений.
- обеспечить целостность и непротиворечивость всей информации в системе,
- избежать дублирования ввода данных,
- улучшить информационное взаимодействие, повысить качество и оперативность работы подсистем.

Таким образом, интегрирование многих факторов является ключевым методом геоинформационных систем. Чем больше информации в системе, тем эффективнее работает система. Причем, эти возможности растут не пропорционально, а экспоненциально.

ГИС интегрирует не только данные, но и передовые информационные технологии. Технологии автоматизированной картографии, технологии баз данных, технологии цифровой обработки результатов позиционирования и дистанционного зондирования Земли, технологии ГИС Web сервисов в открытых средах Интернет и другие не только расширяют возможности ГИС, но и определяют

будущую стратегию развития ГИС в связи с интеграцией и развитием передовых информационных технологий.

1.2.4.2 Анализ

Геоинформационные системы отличаются от других информационных систем тем, что обладают эффективными возможностями анализа пространственных данных и на его основе выполнять пространственное моделирование объектов и явлений. ГИС есть инструмент пространственного анализа. Пространственный анализ называют "сердцем" ГИС.

Развитые ГИС имеют богатый набор средств для выполнения пространственного анализа. Аналитические возможности ГИС позволяют получить ответы на множество пространственных запросов, решить большое количество пространственных задач в разных предметных областях. Множество задач пространственного анализа можно разделить на 5 обобщенных категорий.

1) *Анализ местоположения.*

Этой категории соответствует пространственный запрос: что существует в конкретном месте на поверхности Земли? Чтобы увидеть, где расположен и как выглядит объект, который интересует, используются карты. На карте наглядно представлен характер пространственного распределения объектов, а это позволяет проявить связи между ними и лучше понять исследуемую область. Только увидев местоположения объектов, можно понять некоторые причины пространственных взаимосвязей. Для того, чтобы исследовать закономерности в распределении данных, нужно определенным образом отобразить объекты, которые исследуются, опираясь на значения их характеристик. Например, а) эколог может оценить влияние особенностей рельефа или других факторов на пространственное распределение растительных сообществ, используя картографические данные; б) градостроитель планирует размещение определенного объекта в соответствии с генеральным планом города, планом существующей застройки и планом инженерной инфраструктуры; в) отдел милиции создает карту распределения преступлений разного типа, чтобы видеть, повторяются ли они в отдельных районах.

2) *Удовлетворение пространственных условий.*

Этой категории соответствует пространственный запрос: где удовлетворяются конкретные пространственные условия? Простой запрос о местоположении объекта состоит из одного условия. Для получения ответа достаточно выполнить одну штатную операцию. Сложный запрос о местоположении объекта может включать определенный набор условий. Для получения ответа уже требуется использование ряда операций пространственного анализа. Например, а) где находится площадка для строительства площадью 2 га. в пределах до 200 м. от дороги районного значения, с грунтами несущей способности до I кг. на кв.см.? б) обосновать местоположение торгового, учебного заведения или бизнес-центра с учетом многих, в т.ч. пространственных факторов; в) найти оптимальную трассу трубопровода или путепровода, который проектируется.

3) *Временной анализ.*

Этой категории соответствует запрос: что пространственно изменилось за указанный период? Ответ на этот вопрос представляет собой попытку определить изменения, произошедшие в пространстве и времени, тенденции этих изменений на определенной территории. Например, какова тенденция распространения гриппа в городе, какие новые объекты построены за последний год, каков рост урбанизированных территорий? Сохраняя и сопоставляя карты разных дат, ГИС может выполнять временной анализ.

4) *Выявление структуры.*

Этой категории соответствует пространственный запрос: какие пространственные структуры или распределения существуют? Например, сколько имеется аномалий, не соответствующих нормальному распределению, и где они находятся? Каково распределение населения в городе? Какие участки дороги являются наиболее опасными? Каково распределение стоимости недвижимости на территории? Выделение пространственных структур - это сложный вопрос, требующий применения арсенала мощных средств пространственного анализа.

5) *Оценка различных сценариев.*

Сценарий потенциала является результатом вопросов типа "Что произойдет, если...". Например, что произойдет, если интенсивность дождя будет критической? Каковы издержки, если улицу расширить на 14 м.? Как изменится транспортное сообщение, если убрать трамвай с улицы Пушкинской? В таких случаях пользователь использует модель для прогнозирования и

карты потенциального воздействия. Применение такой модели позволяет построить гипотетическую ситуацию и прогнозировать развитие и следствия социологических и экономических ситуаций, стихийных бедствий и аварий естественного техногенного характера в пространстве и времени.

В последнее время наблюдается заметный рост роли аналитических и моделирующих функций ГИС. Например, система ArcGIS 9.3 (ESRI) включает доступные в освоении модули Spatial Analyst, 3D Analyst, Network Analyst, Geostatistical Analyst.

1.2.4.3 Визуализация

ГИС имеет мощный инструмент визуализации информации. Тематическая информация отображается посредством картографических образов, диаграмм, графиков, оформленных богатым арсеналом изобразительных средств, адаптированных для удобного восприятия информации.

Объекты карты могут быть отображены или напечатаны в любой комбинации и фактически в любом масштабе карты, делая компьютеризированные картографические данные намного более гибкими по сравнению с традиционными бумажными картами.

Большинство людей эффективнее мыслят, если интересующая их проблема представлена в форме картин, чертежей, диаграмм, иллюстраций. Известно, что человек воспринимает 90% информации глазами. Свойства глаза человека позволяют мгновенно оценивать видимую ситуацию в целом. Поэтому говорят, что лучше один график, чем 1000 таблиц. Представьте таблицу для десятка тысяч зданий с 50 характеристиками каждого здания, по которой нужно сделать анализ по ряду показателей. ГИС позволяет трансформировать информацию через привычные с детства картографические образы и мгновенно воспринять ее. ГИС позволяет реализовывать множество запросов и представлять ответы на понятном большинству пользователей языке карты.

ГИС - это средство, помогающее повысить качество принимаемых решений на основании эффективного представления результатов обработки и анализа пространственных данных. Эти способности делают ГИС чрезвычайно полезной для широкого круга людей и организаций, для планирования стратегии и управления инфраструктурой.

1.2.5 ГИС как пересечение областей научных знаний

ГИС – это система, интегрирующая знания. Проблемы разработки, функционирования и использования ГИС находятся на стыке трех областей научных знаний: компьютерные науки, науки о земном пространстве, области ГИС-приложений (Рис. 1.2.4).

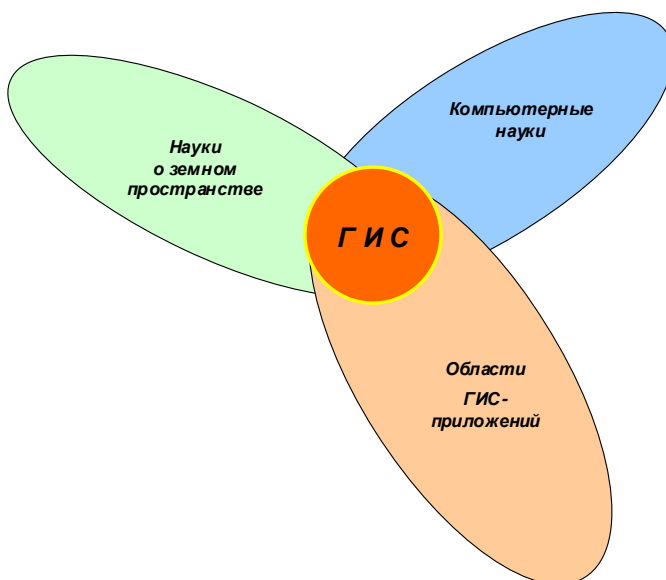


Рис.1.2.4 - ГИС – это пересечение областей научных знаний.

Из такого понимания ГИС следует, что:

- отсутствие в этой совокупности одной любой области научных знаний не приводит к созданию геоинформационной системы в общепринятом смысле.
- пользователи ГИС должны обладать знаниями из трех перечисленных областей.

Табл. 1.2.1 - Базовые области научных знаний ГИС.

Компьютерные науки	Науки о земном пространстве	Области ГИС-приложений
Информатика и программирование, математическое моделирование, операционные системы, текстовые и графические редакторы, электронные таблицы, базы данных, информационные сети, обработка изображений, ...	Геодезия, картография, география, аэрокосмическая съемка, дистанционное зондирование Земли, глобальные системы позиционирования, ...	Управление территорией, градостроительство и архитектура, инженерная инфраструктура, управление недвижимостью, транспорт и логистика, экология, природные ресурсы, демографические исследования, оборона, сельское хозяйство, Правонарушения, чрезвычайные ситуации, ... (около 80 дисциплин)

1.2.6 Примеры применения ГИС

Спектр применений ГИС чрезвычайно широк. ESRI перечисляет около 80 областей использования ГИС. В перечислении ГИС приложений можно выделить базовые типы распространенных задач [14, с.121-134].

- 1) *Задачи учетно-инвентаризационного типа.* В них акцент делается на данных, измерениях и оценке "степени похожести" (задачи земельного кадастра, подсчеты запасов природных ресурсов, управление распределенной производственной инфраструктурой). Это самый распространенный тип приложений ГИС. Для приложений такого типа характерна работа с большим числом

географических объектов и высокая детальность изучения территорий.

- 2) *Задачи и цели деятельности.* Здесь основные задачи - планирование развития, выбор маршрутов и управление перевозками. Специфика приложений этого типа связана с нетрадиционными постановками оптимизационных задач, заданными на структурах допустимых путей.
- 3) *Моделирование и сложные методы анализа данных.* Типичная задача - предсказание наводнений через анализ водосбора и водостока на заданном рельефе.

В качестве примеров рассмотрим некоторые области, где применение ГИС стало уже традиционным.

- *Управление и планирование развития территорий.* Эта область основана на предполагаемом поведении различных социальных групп, определяющих общественные потребности и возможности и имеющих заданное или предполагаемое размещение и динамику в рамках заданной территории.
- *Градостроительство и архитектура.* Проектирование, инженерные изыскания, планировка в градостроительстве, архитектуре. Это типичная работа городских служб, обеспечивающих нормальное развитие подведомственной территории.
- *Инженерная инфраструктура.* Инвентаризация, учет, планирование размещения объектов распределенной производственной инфраструктуры (водоснабжения, водоотведения, теплоснабжения, газоснабжения, электроснабжения) и управление ими, оценка состояния и принятие решений при ремонтных или аварийных ситуациях.
- *Управление земельными ресурсами, земельные кадастры.* Область характерна своей чисто географической ориентацией. Типичные задачи здесь - составление кадастров, классификационных карт, определение границ участков и площадей и т. д.
- *Управление природными ресурсами и природоохранная деятельность.* Здесь типичными проблемами являются определение текущих состояний и запасов наблюдаемых ресурсов, моделирование процессов в природной среде и построение обоснований принимаемых решений по управлению средствами, изменяющими природные ресурсы.
- *Геология, минерально-сырьевые ресурсы, горно-добывающая промышленность.* Специфика таких проблем состоит в том, что требуется рассчитать запасы полезных ископаемых в некоторой

области по результатам определения в отдельных точках (разведочное бурение, пробные шурфы и т.д.) при известной модели процесса образования месторождения.

- *Планирование и управление перевозками (логистика).* Заданы (на карте) пункты со своими характеристиками, в которых хранятся грузы, и пункты со своими характеристиками, которые ожидают нужные им грузы; средства перевозки со своими характеристиками, позицией, состоянием и специализацией; сеть дорог со своими характеристиками (средняя скорость, ремонты, объезды, пробки, границы, таможенные пункты и т.д.). Требуется составить план перевозок и корректировать его по мере возникновения непредвиденных ситуаций.
- *Наземное, аэро- и гидронавигационное картографирование и управление наземным, воздушным и водным транспортом.* Традиционные области с понятными проблемами. Особое место здесь занимают проблемы управления движущимися объектами при условии выполнения заданной системы отношений между ними и неподвижными объектами.
- *Маркетинг и анализ рынка.* Определение тенденций развития ситуации, оценка влияния различных топологических свойств - близости, пересечений и совмещений различных ареалов на их взаимодействие, учет разных условий, определенных на объектах с заданными позициями, потребностями и возможностями их развития.
- *Сельское хозяйство.* Подсчет запасов ресурсов по ряду точечных измерений, планирование перевозок, взаимодействие динамически изменяющихся ареалов, категоризация и выделение "похожести" пространственных объектов, точное земледелие.
- *Чрезвычайные ситуации.* Учет потенциально опасных объектов, моделирование последствий в чрезвычайных ситуациях.
- *Службы быстрого реагирования.* Общественная безопасность, пожаротушение, скорая медицинская помощь.

Дополнительно к перечисленным областям следует также отметить государственное, региональное, муниципальное управление, демография, зонинг, регистрация собственности, налогообложение, банковское дело, сферу обслуживания, торговля, бизнес, социальное страхование, мониторинг природной среды, экология, оценку возможных воздействий на природу, водные ресурсы, гидрология, почвоведение, лесное хозяйство, область военного применения...

1.2.7 Значение ГИС для управления городом

Одна из областей эффективного применения ГИС стала область местного самоуправления. Понимание роли геоинформационного обеспечения управления привело к тому, что ГИС стали де-факто мощным общепринятым инструментом в области государственного и муниципального управления во многих странах. Только в Европе более 100000 муниципалитетов используют ГИС для городского управления. ГИС рассматривается как инструмент для принятия решений в области управления городским хозяйством, как средство планирования и выработки стратегии развития территорий городов.

Способность ГИС объединять разноплановую информацию разных структур городского управления в единый информационный комплекс позволяет обеспечить взаимодействие управляющих структур города. ГИС — это современная технология для совершенствования городского управления.

Внедрение муниципальных ГИС означает переход на новый, более высокий качественный уровень управления, планирования, проектирования, эксплуатации сложных систем городского хозяйства. Этот уровень характеризуется следующими свойствами:

- 1) управление городом в реальном времени;
- 2) минимизация и исключение дублирования функций исполнительных органов на основе системного интегрирования городских данных;
- 3) обеспечение взаимодействия управляющих структур города на основе создания и использования единой базы городских данных;
- 4) снижение расходов за счет экономии времени при выполнении функций с большим эффектом;
- 5) улучшение производительности труда;
- 6) повышение сбора налогов.

Поэтому, внедрение ГИС в городское управление является актуальной проблемой всех городов Украины.

1.2.8 Историческая справка о геоинформационных системах

Анализ источников по ГИС, проведенный Кошкарёвым А.В. и Тикуновым В.С. и др. [15], [16], [17] показывает, что истоки рождения геоинформатики следует искать в работах коллективов, сформулировавших первые задачи и подходы к построению информационных систем, ориентированных на обработку

пространственных данных в Канаде и Швеции - двух странах, приоритет которых в этой области абсолютно бесспорен.

Канадские работы были связаны с созданием в 1963-1971 гг. Канадской ГИС (Canadian Geographical Information System - CGIS) под руководством Роджера Томплинсона. Ставшая одним из примеров крупной универсальной региональной ГИС национального уровня; CGIS может считаться классикой.

Работы шведской школы геоинформатики концентрировались вокруг ГИС земельно-учетной специализации, в частности Шведского земельного банка данных, предназначенного для автоматизации учета земельных участков (землевладения) и недвижимости.

Период 1960-х – ранних 1970-х годов.

Это период исследований принципиальных возможностей, пограничных областей знаний и технологий, наработки эмпирического опыта, первых крупных проектов и теоретических работ. Это период:

- появления цифрователей, плоттеров, графических дисплеев и других периферийных устройств;
- создания программных алгоритмов и процедур графического отображения информации на дисплеях и с помощью плоттеров;
- создания формальных методов пространственного анализа;
- создания программных средств управления базами данных.

Ранние ГИС значительно отличались от того, что понимается под ГИС сегодня. Их отличала ориентация на задачи инвентаризации земельных ресурсов, земельного кадастра и учета в интересах совершенствования системы налогообложения, решаемые путем автоматизации земельно-учетного документооборота. Основная функция ГИС состояла в вводе в машинную среду первичных учетных документов для хранения и регулярного обновления данных, включая агрегацию данных и составление итоговых отчетов статистических табличных документов.

Инвентаризационные задачи путем массового цифрования карт решались первоначально именно в Канадской ГИС. В ее основу были заложены фундаментальные принципы, которые позволили выйти в сферы не только узко профильных задач, но и более универсальных интересов. Первый шаг, который вывел ГИС из области баз данных общего назначения, заключался во введении в число атрибутов операционных объектов (земельных участков, строений, физических и юридических лиц, ареалов использования земель) признака пространства, в какой бы форме местоуказания (в координатах, в иерархии административной принадлежности, в терминах

принадлежности к ячейкам регулярных сетей членения территории) он ни выразался. Достаточно революционным являлось уже указание координат центроидов объектов. Создатели ГИС Канады внесли следующие новшества:

- использование сканирования для автоматизации процесса ввода геоданных;
- расчленение картографической информации на тематические слои и разработка концептуального решения о "таблицах атрибутивных данных", что позволило разделить файлы плановой (геометрической) геоинформации о местоположении объектов и файлы, содержащие тематическую (содержательную) информацию об этих объектах;
- функции и алгоритмы оверлейных операций с полигонами, подсчет площадей и других картометрических показателей.

В этот период сформировалось понятие пространственных объектов, описываемых позиционными и непозиционными атрибутами. Оформились две альтернативные линии представления - растровые и векторные, включая топологические линейно-узловые представления. Чуть позже создана технология массового цифрования карт - основного источника данных в Канадской ГИС. Поставлены и решены задачи, образующие ядро геоинформационных технологий: наложения (оверлей) разноименных слоев, генерация буферных зон, полигонов Тиссена и иные операции манипулирования пространственными данными, включая определения принадлежности точки полигону, операции вычислительной геометрии вообще.

Большое воздействие на развитие ГИС оказала Гарвардская лаборатория компьютерной графики и пространственного анализа (Harvard Laboratory for Computer Graphics & Spatial Analysis) Массачусетского технологического института (руководитель - Howard Fisher). Программное обеспечение Гарвардской лаборатории широко распространялось и помогло создать базу для развития многих ГИС-приложений. В этой лаборатории были заложены основы картографической алгебры, создано семейство растровых программных средств Map Analysis Package - MAP, PMAP, aMAP. Наиболее известными программными продуктами Гарвардской лаборатории являются: SYMAP (система многоцелевого картографирования), CALFORM (программа вывода картографического изображения на плоттер), SYMVU (просмотр перспективных трехмерных изображений), ODYSSEY (предшественник знаменитого ARC/INFO).

Функциональная ограниченность ГИС первого поколения имела чисто технические причины. Незрелость периферийных устройств, пакетный режим обработки данных (без дисплея), критичность вычислительных ресурсов и времени вычисления задач. Ядро ГИС было сформировано в конце 60-х, определив облик ГИС первого поколения.

Период начала 1970-х – начала 1980-х годов.

Для 70-х годов характерна разработка взаимодействия методов и средств геинформатики с цифровыми методами картографирования и автоматизированной картографией. ГИС в современном их понимании развивались на базе информационно-поисковых систем, позднее приобретая функции картографических банков данных с возможностью моделирования и анализа данных. Большинство ГИС этого периода включает в свои задачи создания карт и используют картографический материал как источник данных. К этому периоду относится быстрый прогресс геоинформационных технологий в США.

К началу 1970-х годов в США сформировалось мнение о необходимости использования ГИС-технологий для обработки и представления данных переписи населения. Потребовалась методика, обеспечивающая корректную географическую "привязку" данных переписи. Основной проблемой стала необходимость конвертирования адресов проживания населения в географические координаты так, чтобы результаты переписи можно было бы оформлять в виде карт по территориальным участкам и зонам Национальной переписи. Был разработан специальный формат представления картографических данных, для которого были определены прямоугольные координаты перекрестков, разбивающих улицы всех населенных пунктов США на отдельные сегменты. Алгоритмы обработки и представления картографических данных были заимствованы у разработчиков ГИС Канады и Гарвардской лаборатории и оформлены в виде программы POLYVRT, осуществляющей конвертирование адресов проживания в соответствующие координаты, описывающие графические сегменты улиц.

Таким образом, в этой разработке впервые был широко использован топологический подход к организации управления географической информацией, содержащий математический способ описания пространственных взаимосвязей между объектами.

Период 1980-х годов.

80-е годы отличает чрезвычайный динамизм развития ГИС. Широкий рынок разнообразных программных средств, развитие настольных ГИС, расширение области их применения за счет интеграции с базами непространственных данных, появление сетевых приложений, значительного числа непрофессиональных пользователей, систем, поддерживающих индивидуальные наборы данных на отдельных компьютерах, открывают путь системам с распределенными базами геоданных.

К середине 80-х их число ГИС-пакетов приближается к 500. Разработка коммерческих ГИС связана в большей степени с возможностями ЭВМ, а также с ПК. Создание ГИС стало основываться не на уникальных программных и аппаратных средствах, доступных только хорошо финансируемым организациям (типа министерства обороны), но и для небольших компаний, образовательных и муниципальных учреждений, и даже для частных лиц.

Популярность в мире получил пакет анализа растровых данных MAP (Map Analysis Package), который реализовал алгоритмы картографической алгебры, основы которой были разработаны С.Д. Томлином, США в 1983 г. В этот период были разработаны теоретические основы геостатистики (Ж. Матерон, Франция), векторная топологическая структура пространственных данных (DIME-структура, США), технологии графического изображения трехмерных поверхностей и др.

В 1981 г. был разработан знаменитый программный продукт ARC/INFO Институтом исследования систем окружающей среды (Environmental System Research Institute, ESRI Inc.) Он был наиболее успешным воплощением идей о раздельном внутреннем представлении геометрической и атрибутивной информации. Для хранения и работы с атрибутивной информацией в виде таблиц INFO был успешно применен формат стандартной реляционной системы управления базами данных, а для хранения и работы с графическими объектами в виде дуг (ARC) было разработано специальное программное обеспечение. ARC/INFO стал первым программным пакетом ГИС, который эффективно использовал пользовательские качества персональных компьютеров, в то же время он доступен для разных технических платформ и операционных сред.

В эти годы разрабатываются: пакет MapInfo фирмы Mapping Information Systems Corp. (1987), пакет IDRISI в Университете Кларка (1987), пакет Modular GIS Environment (MGE) фирмы Intergraph, Corp.

(1988), CORINE - Геоинформационная система страны Европейского содружества (1985 г.) и GRID - Глобальный ресурсный информационный банк данных (1987 г.). В Европе основные шаги в разработке и внедрению ГИС-технологий были сделаны в Швеции, Франции, Нидерландах, Великобритании и Западной Германии Осваиваются принципиально новые источники данных для ГИС: данные дистанционного зондирования, включая материалы спутников серии Landsat, Spot.

Период 1990-е годы - настоящее время.

Этот период характеризуется новым отношением к пользователям, повышенной конкуренцией среди коммерческих производителей геоинформационных технологий и услуг, возросшей потребностью в геоданных, началом формирования мировой геоинформационной инфраструктуры.

Разработчики геоинформационного программного продукта GRASS для рабочих станций, созданного американскими военными специалистами (Army Corps of Engineers) для задач планирования природопользования и землеустройства открыли его для бесплатного пользования. В результате пользователи и программисты могут создавать собственные приложения, интегрируя GRASS с другими программными продуктами.

В этот период Институт исследования систем окружающей среды (ESRI Inc.) разрабатывает программное обеспечение ГИС, которое вывело его в мировые лидеры: 1995 г. - ArcView v.1.0, 2001 г. - ArcGIS v. 8.1, 2004 г. - ArcGIS v. 9.0, 2005г. - ArcGIS v. 9.1, 2006 ArcGIS v. 9.2, 2008 г. - ArcGIS v. 9.3 - флагман программного обеспечения ГИС.

В Советском Союзе исследования в отрасли геоинформационных технологий были начаты в восьмидесятые годы. Осуществлялись исследования в области автоматизации картографирования, пространственного анализа, картографо - математического моделирования, (О. М. Берлянт, Н. Л. Беручишвили, В. Т. Жуков, Д. В. Лисицкий, П. В. Петров, С. М. Сербенюк, Ю. Г. Симонов, В. С. Тикунов, И. Г. Черванев, В. А. Червяков и др.), теоретического обоснования геоинформационных систем (Н. Л. Беручишвили, И. В. Гармиз, В. С. Давидчук, В. П. Каракин, А. В. Кошкарев, В. Г. Линник, М. В. Панасюк, А. М. Трофимов и др.).

Первые программные ГИС-пакеты на территории бывшего Советского Союза были разработаны в 90-е годы. Среди них известными являются пакет GeoDraw/GeoGraph Института географии Российской академии наук, пакет "Панорама" Топографической

службы Вооруженных сил Российской Федерации и другие. Однако большая часть программных продуктов ГИС представлена продукцией западных фирм - ESRI, Intergraph, MapInfo, Autodesk и др.

В Украине геоинформационные технологии получили развитие в середине 90-х годов [18] с разработкой и внедрением ГИС-проектов в отрасли управления земельными ресурсами, мониторинга экологического состояния среды, управления территорией, геологии, водных ресурсов, МЧС и других.

Состояние применения геоинформационных технологий в Украине характеризуют следующие факторы:

- формирование в государственных учреждениях и организациях группы специалистов, которые активно работают в направлении применения ГИС в разных сферах деятельности;
- создание предприятий, которые специализируются в разработке или использовании геоинформационных технологий, в частности: государственного научно-производственного предприятия "Геосистема", в котором разработано программное обеспечение Delta/Digitals (г. Винница) и научно-производственного центра "Геодескартинформатика" (г. Киев); коммерческих компаний "ЕСОММ", "Интеллектуальная система, Гео", "Институт передовых технологий", ГЕОКАД, "Аркада", "Геоника" (г. Киев); СПАЭРО+ (г. Харьков), "Высокие технологии" (г. Одесса) и др.;
- внедрение цифрового картографирования в системе Укргеодезкартографии;
- открытие подготовки специалистов в области геоинформационных систем и технологий в 8 вузах Украины.

В заключение этой темы следует обратить внимание на то, что обстоятельные исследования истории ГИС приведены в источнике [17].

1.2.9 Контрольные вопросы и задания для самостоятельной работы

- 1) Дайте определение ГИС. В чем состоит отличие ГИС от других информационных систем?
- 2) Приведите общую характеристику компонентов ГИС.
- 3) Приведите общую характеристику отличительных функций ГИС от функций других информационных систем?
- 4) Какие области научных знаний интегрирует ГИС?
- 5) Охарактеризуйте историю развития ГИС.

Раздел 1.3

ГИС-ПАРАДИГМА

Относительно короткий период бурного развития ГИС естественно сопровождается в ряде случаев неодинаковым представлением предметной области геоинформационных систем и технологий. Различия во взглядах, неоднозначное трактование базовых понятий и определений говорят не только о неодинаковом их понимании, но и о сложности выработки единых подходов.

"Парадигма (от греческого – пример, образец) – теория (или модель постановки проблем), принятая в качестве образца решения исследовательских задач" [19]. Здесь слово парадигма используется для определения модели или каркаса, системы устоявшихся взглядов на понимание области геоинформационных систем и технологий. ГИС-парадигма – это концептуальное основание для использования географической информации, которая обеспечивает общий базис соотнесения или фокусирования для принципов управления данными, технологий и применений.

Область геоинформационных систем и технологий требует фиксирования и представления устоявшихся в мире базовых понятий и определений, их однозначного понимания. В первую очередь требуются ответы на вопросы, что есть географическая информация, географические данные, геоинформационные технологии, геоинформационные системы, геоинформационная наука, геоинформационное образование, геоинформатика? Здесь представлена попытка изложения единой системы базовых понятий и определений в области геоинформационных систем и технологий.

1.3.1 Что есть географическая информация, географические данные?

В Национальном стандарте Украины ДСТУ ISO 19101:2002(Е) дается такое определение: *"Географічна інформація (Geographic Information) - Інформація про об'єкти та явища, які безпосередньо або опосередковано пов'язані з певним місцеположенням відносно Землі"* [19, (4.16)]. Иными словами, *географическая информация* – это информация об объектах, системах объектов, явлениях и процессах реального мира, которые имеют или могут иметь пространственную привязку в реальном пространстве Земли.

В работе [13] приводится такое пояснение: географическая информация – это информация об участках поверхности Земли, знание относительно того, где находится нечто, знание относительно того, что является в данном месте. Географическая информация имеет отношение к Земле, ее двумерной поверхности, ее трехмерному пространству.

В формализованном виде географическая информация представляется *географическими данными*. Географические данные представляют собой единство двух компонентов – геопространственного и тематического.

- Геопространственный компонент является определяющим понятия "географическая информация". Он характеризует локальные свойства земного пространства – местоположение, форму, размеры, пространственные отношения макрообъектов Земли. Пространственный компонент имеет два ключевых аспекта: абсолютное местоположение, которое определяется системой координат, и топологические отношения к другим сущностям.
- Тематический компонент: это переменные или атрибуты, которые могут быть изучены, рассматривая тематический аспект, аспект местоположения или оба аспекта.

Термин "геопространственные данные" (Geospatial Data) введен в официальный язык в 1994 г. в распоряжении президента У. Клинтона [10].

1.3.2 Что есть Геоинформационные технологии?

Геоинформационные технологии (Geographic Information Technologies) – это совокупность методов и приемов для сбора и обработки географической информации. В работах [13], [20], [21] выделены три главных типа геоинформационных технологий:

- *Технологии собственно геоинформационных систем* (Geographic Information Technologies) – это технологии ввода, интегрирования, хранения, обработки, анализа, моделирования и визуализации географической информации.
- *Технологии дистанционного зондирования Земли* (Remote Sensing - RS) – это технологии получения информации о поверхности Земли и среды посредством использования орбитальных спутников Земли. Сигналы, переданные к приемникам на Земле, преобразуются в цифровые изображения для изучения, обработки и интерпретации.

- *Технологии позиционирования* (Global Positioning System – GPS) – это технологии определения местоположения на Земле, обработка данных интегрированных в среду геоинформационных систем средствами глобальных навигационных систем, комбинированными средствами глобальных навигационных систем и средствами электронных геодезических измерений.
- Обработка данных дистанционного зондирования Земли и данных позиционирования выполняется интегрированными в ГИС средствами, что дает основание относить эти технологии к геоинформационным технологиям.

1.3.3 Что есть Географическая информационная система?

Среди многих определений Географической информационной системы (*Geographic Information System - GIS*) базовыми определениями, построенными весьма авторитетными специалистами в ГИС-сообществе, представляются следующие определения:

- 1) "Are two distinct meanings of the question "is this a GIS?":
 - a) GIS is a real application, including the hardware, data, software and people needed to solve a problem (a GIS application).
 - b) GIS is a type of software sold by a software developer (compare Microsoft Word).
 - c) We will focus on #1 first" [12, Michael F. Goodchild]
- 2) "Like the field of geography, the term Geographic Information System (GIS) is hard to define. It represents the integration of many subject areas. Accordingly there is no absolutely agreed upon definition of a GIS" [11, Michael N. DeMers].
- 3) "A broadly accepted definition of GIS is: a GIS is a system of hardware, software and procedures to facilitate the management, manipulation, analysis, modelling, representation and display of georeferenced data to solve complex problems regarding planning and management of resources" [13, National Centre of Geographic Information and Analysis]
- 4) "Географічна інформаційна система - інформаційна система, що опрацьовує інформацію про об'єкти та явища, які пов'язані з певним місцеположенням відносно Землі". [19, 4.18]

На основании рассмотрения этих и ряда других известных определений, выделения в них общих компонентов в предыдущем разделе представлено следующее системное определение, которое соответствует наиболее распространенному пониманию этой области:

Географическая информационная система, ГИС (Geographic Information System, GIS) – это система, которая:

во-первых, представляет собой совокупность взаимодействующих пяти компонентов - компьютерных средств, программного обеспечения, географических данных, регламента и пользователей;

во-вторых, выполняет функции ввода, интегрирования, хранения, обработки, анализа, моделирования и визуализации географической информации.

1.3.4 Что есть геоинформационная наука?

Работы Консорциума 65 университетов для географической информационной науки (University Consortium for Geographic Information Science, UCGIS) [22], Девида М.Марка [20], Национального центра географической информации и анализа [13], Университета штата Калифорния [23], В.П.Савиных [24]. позволяют сделать следующее определение геоинформационной науки.

Геоинформационная наука (Geographic Information Science, GI Science) – это область исследований, направленных на развитие и использование теории, методов, технологий и данных для понимания географических процессов, отношений и структур, пространственно-временных изменений на Земле в контексте географических информационных систем.

Геоинформационную науку определяют:

- *объект изучения* - географические системы, отношения, распределения и структуры, пространственно-временные изменения объектов и явлений реального мира;
- *предмет изучения* - пространственные теории, методы и технологии, используемые в контексте географических информационных систем;
- *методология изучения* - методология междисциплинарного исследования, основанная на использовании методов информатики и геоинформационного моделирования объектов и явлений реального мира.

Области исследования и разработки геоинформационной науки, согласно [13], следующие.

- 1) Проблемы представления систем, процессов и явлений Земного пространства:
 - поверхность Земли - бесконечный континуум,

- какие критерии могут использоваться для выбора представления?
- 2) Проблемы оценки представления:
- как измерять точность представления,
 - как измерять то, что является отсутствующим, то, что является неопределенностью,
 - как выражать их способами, которые являются значимыми для потребителя, как описывать, как визуализировать, как моделировать?
- 3) Проблемы относительно отношений между представлением и пользователем:
- как люди думают относительно мира,
 - как компьютерные представления могут сделать больше, чем мышление Человека,
 - как люди рассуждают, узнают, связывают относительно географического мира,
 - какой выход от ГИС можно сделать более понятным?
- 4) Проблемы определения неявных структур, распределений, взаимосвязей:
- проблемы относительно отображения географических данных,
 - как методы отображения эффективны для истолкования географических данных,
 - как наука картография может быть расширена посредством использования преимуществ цифровой среды,
 - какие базисные свойства отображения определяет его успех?
- 5) Проблемы относительно аналитических инструментальных средств:
- какова природа человеческой пространственной интуиции, и как она может быть расширена инструментальными средствами ГИС,
 - какие методы анализа необходимы, чтобы поддержать определенные типы решений, сделанные с использованием ГИС,
 - как методы анализа могут быть представлены, чтобы пользователи могли выбирать эффективно между ними?
- 6) Проблемы относительно моделей данных и структур:
- как сохранять представление эффективно,
 - как быстро отыскивать информацию через соответствующую индексацию,
 - как достичь взаимодействия между системами?

1.3.5 Что есть Геоинформационное образование?

Геоинформационное образование, ГИС-образование (GIS education) определяется совокупностью знаний геоинформационной науки и технологий.

Совокупность знаний ГИСиТ, детально представленная в работе [25] Консорциума университетов для географической информационной науки (UCGIS), содержит следующие компоненты.

- 1) Аналитические методы:
 - академические и аналитические начала;
 - операции запросов и язык запросов;
 - геометрические измерения;
 - базисные аналитические операции;
 - базисные аналитические методы;
 - анализ поверхностей;
 - пространственные статистики;
 - геостатистики;
 - пространственная регрессия и эконометрия;
 - сбор данных;
 - сетевой анализ;
 - оптимизация и моделирование местоположения-распределения.
- 2) Концептуальные основания:
 - философские основания;
 - когнитивные и социальные основания;
 - домены географической информации;
 - элементы географической информации;
 - отношения;
 - неоднородности в географической информации.
- 3) Картография и визуализация:
 - история и тенденции;
 - рассмотрение данных;
 - принципы картографического дизайна;
 - техника графического представления;
 - картографическая продукция;
 - использование и оценки карт.
- 4) Аспекты проектирования:
 - область системного проектирования ГИСиТ;
 - определение проекта;
 - ресурсы планирования;
 - разработка базы данных;

- планирование анализа;
 - планирование приложений;
 - системы внедрения.
- 5) Моделирование данных:
- основные структуры хранения и поиска;
 - системы управления базами данных;
 - мозаичные модели данных;
 - векторные и объектные модели данных;
 - 3D-моделирование не точных и временных явлений.
- 6) Обработка данных:
- представление трансформаций;
 - генерализация и агрегирование;
 - управление транзакциями.
- 7) Геовычисления:
- появление геовычислений;
 - аспекты вычислений и нейровычисления;
 - клеточные автоматы (КА);
 - эвристики;
 - генетические алгоритмы;
 - агентные модели;
 - имитационное моделирование;
 - неопределенность;
 - нечеткие множества.
- 8) Геопространственные данные:
- геометрия Земли;
 - системы разделения земли;
 - геореференционные системы;
 - даты;
 - картографические проекции;
 - качество данных;
 - геодезическая съемка и GPS;
 - оцифровка;
 - сбор полевых данных;
 - аэрофотоизображения и фотограмметрия;
 - дистанционное зондирование со спутника и с борта судна;
 - метаданные, стандарты и инфраструктуры.
- 9) ГИС и общество:
- правовые аспекты;
 - экономические аспекты;
 - использование геопространственной информации;
 - геопространственная информация как собственность;

- распространение геопространственной информации;
- этические аспекты;
- критические ГИС.

10) Организационные и институциональные аспекты:

- начала ГИСиТ;
- управление геоинформационными системами и инфраструктурой;
- организационные структуры и процедуры;
- темы трудовых ресурсов ГИСиТ;
- институциональные и между-институциональные аспекты;
- организация координирования.

Более детально совокупность знаний представлена в работе [25]. Это исчерпывающий перечень компонентов знаний, который должен определять ГИС-образование и лежать в основе определения содержания ГИС-образования.

В зависимости от объема знаний GIS Educator ESRI различает следующие *уровни* ГИС-образования, представленные на Рис. 1.3.1.

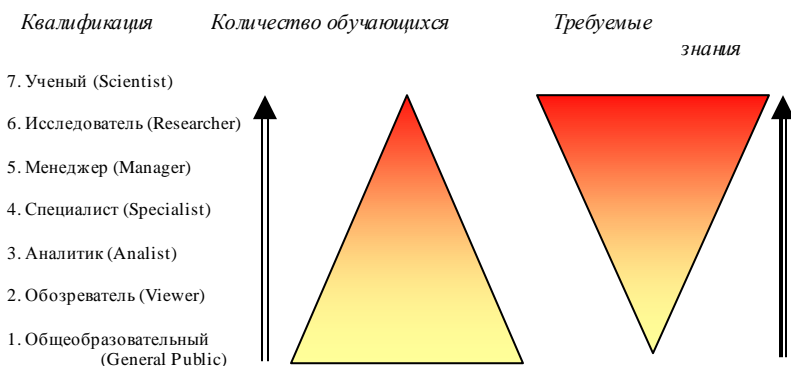


Рис. 1.3.1 – Уровни ГИС образования

1.3.6 Что есть геоинформатика?

В работах Кошкарева А.В. и Тикунова В.С. [15, 16], Берлянта А.М. [26] дано следующее определение геоинформатики:

Геоинформатика (Geoinformatics) - научно-технический комплекс, объединяющий одноименную отрасль научного знания,

технологии и прикладную (производственную) деятельность, которые связаны со сбором, хранением, обработкой и отображением пространственных (географических) данных, а также с проектированием, созданием и эксплуатацией ГИС.

1.3.7 Контрольные вопросы и задания для самостоятельной работы

- 1) Что есть географическая информация, географические данные?
- 2) Приведите определения понятий Географические информационные системы, Геоинформационные технологии, Геоинформационная наука.
- 3) Каково должно быть содержание Геоинформационного образования?
- 4) В чем состоит различие между понятиями Географические информационные системы и Геоинформатика?

Часть 2

**КОМПЬЮТЕРНЫЕ МОДЕЛИ
ГЕОГРАФИЧЕСКИХ ОБЪЕКТОВ**

Раздел 2.1

ГЕОГРАФИЧЕСКИЕ ОБЪЕКТЫ

2.1.1 Пространство и время

Выделение объекта и предмета познания любой науки опирается на систему наиболее общих понятий и законов. Для геоинформационной науки фундаментом решения этой проблемы может быть философская категория пространства-времени. Пространство и время существуют не отдельно, а вместе, образуя непрерывное многообразие - пространственно-временной континуум. Пространство является единой формой существования материи [3, с. 541-542].

- Пространство - это форма сосуществования материальных вещей и явлений.
- Время - это форма выражения сменяемости состояния материальных объектов и явлений.

Таким образом, каждая вещь, каждый материальный объект находятся не в пространстве и времени, как обычно говорят, а является частью, "куском" некоторой области пространства-времени.

Категория пространства характеризуется определенными свойствами.

- Свойства общие для пространства и времени: абсолютность, безграничность, бесконечность, связность, непрерывность, объективность.
- Свойства специфические для пространства: а) общие - протяженность, не направленность, трехмерность, структурность; б) локальные - форма, размеры, местоположение, распределение вещества, распределение поля, симметрия и асимметрия.
- Свойства специфические для времени: а) общие - длительность, направленность, одномерность, последовательность, б) локальные - периоды, моменты, одновременность, ритм, скорость, асимметрия.

Различают также свойства метрические (связанные с измерениями) и топологические (связность, непрерывность).

2.1.2 Определение географических объектов

Географические информационные системы моделируют реальное пространство земной поверхности, которое называют географическим пространством. Географическое пространство представляет собой непрерывное разнообразие частей реального земного пространства - континуум различных географических объектов. Между объектами формируется сложная система отношений, развивающихся во времени.

Географические объекты – это сущности географического пространства (объекты, явления, процессы макромира), которые имеют или могут иметь пространственную привязку (локализацию) в реальном земном пространстве. Согласно ISO 19101 географический объект (Feature) – это абстракция реального мира. Географический объект – это средство моделирования объекта реального мира. Термин "географический" ассоциируется здесь с характеристикой пространственности.

Примеры географических объектов: здание, человек, плотность населения, дерево, лес, область заражения, промышленная зона, граница, дорожно-транспортное происшествие, зона шума, почвы, доход, микрорайон, инженерные коммуникации, дорога, река, административно-территориальная единица, земельный участок...

2.1.3 Виды географических объектов

По способу декомпозиции земного пространства различают следующие виды географических объектов: дискретные географические объекты, непрерывные явления, объекты, обобщенные по площади.

Дискретные географические объекты – это отдельные ограниченные макротела реального земного пространства. В любом месте земного пространства дискретные географические объекты могут находиться или отсутствовать. Например, колодцы, ДТП, дороги, трубопроводы, здания, кварталы, зоны...

Непрерывные явления (поля) характеризуют территорию в целом, а не отдельные объекты. Например: поверхности, осадки, температура могут измеряться в любом месте территории и характеризовать ее в целом. Непрерывность явлений проявляется в том, что невозможно указать промежутки на площади распространения явлений, в которых бы они отсутствовали. Несмотря на то, что данные меняются непрерывно, границами указывают дискретное изменение величины в

определенных пределах (например, типы почв). Непрерывные объекты заполняют всю моделируемую поверхность, "пронизывая" друг друга, их можно трактовать как свойства пространства или самой моделируемой поверхности.

Объекты, обобщенные по площади, отображают обобщенную характеристику или концентрацию отдельных объектов в пределах данной области. Статистические показатели обобщаются для определенной территории. Обобщенные показатели характерны для административных районов, почтовых отделений, подрайонов милиции...Пример объектов, обобщенных по площади: количество домов в округе, плотность улично-дорожной сети, население жилых кварталов.

По структуре географические объекты подразделяются на элементарные (простые), составные и сложные.

- Элементарный объект не имеет составных частей (например, отдельное здание).
- Составной объект образуется группой других объектов с определенным (направленным) порядком их следования при образовании определяемого объекта (например, отдельное здание, состоящее из частей).
- Сложный объект образуется группой других объектов (элементарных, составных, сложных), порядок следования которых при образовании определяемого объекта не фиксирован (например, ансамбль зданий).

По форме существования географические объекты подразделяются на материальные (реальные) и абстрактные (виртуальные).

- Реально существующие географические объекты, явления и события могут восприниматься либо непосредственно, с помощью органов чувств, либо опосредованно, с применением приборов
- Виртуальные объекты могут не существовать в реальности, но быть отображаемыми. К ним относятся объекты: существовавшие в прошлом, существование которых предполагается в будущем и воображаемые. Воображаемыми объектами являются границы, красные линии, горизонталы и т. п.

2.1.4 Способы локализации географических объектов

Локализация – это средство указания местоположения на Земле. Локализация возможна посредством координат, указания адреса, указания района.

Локализация посредством координат является наиболее точной. Могут быть использованы географические координаты (широта, долгота) или декартовы координаты проецирования на плоскость (абсцисса, ордината).

Табл. 2.1.1 - Пример локализации городов посредством географических координат

Название города	Широта	Долгота	Количество жителей
Лондон	51 с	0	5000
Цюрих	47 с	8 в	300
Лейпциг	51 с	12 в	700
Мадрид	40 с	4 в	3000
Харьков	49 с	36 в	1500

Локализация посредством указания района широко распространена в обыденной жизни. Для однозначной локализации район должен иметь определенные границы. Районы с размытыми границами, например, исторически сложившиеся поселки или территории вблизи объектов, не имеют точной привязки.

Табл. 2.1.2 - Пример локализации купли-продажи квартир указанием района.

Название района	Комнат	Площадь	Этаж	Этажей	Цена
Сев. Салтовка	1	17,5	6	9	5400
Алексеевка	1	17,5	10	12	6500
Павлово поле	1	17,5	5	5	7200
Район ХТЗ	1	17,5	3	12	6000
Район ХТЗ	2	27	3	4	8500
Павлово поле	2	30	3	5	9500
Сев. Салтовка	2	28	9	9	8050
Новые дома	2	31	2	5	8000

Локализация посредством указания адреса наиболее распространенная текстовая форма.

Табл. 2.1.3 - Пример локализации объекта посредством указания адреса объекта.

Объект	Адрес
Горисполком	Пл. Конституции, 7
Академия	Ул. Революции, 12
Иванов И.И.	Ул. Пушкинская, 36, кв.18

2.1.5 Контрольные вопросы и задания для самостоятельной работы

- 1) Что такое географические объекты?
- 2) Какие различают виды географических объектов?
- 3) Приведите характеристики способов локализации географических объектов?

Раздел 2.2

ГЕОГРАФИЧЕСКИЕ ДАННЫЕ

2.2.1 Понятия "информация" и "данные"

Информация

Информация как концепция имеет многообразие значений от ежедневного использования до технического применения. В большинстве случаев концепция информации тесно связана с понятиями ограничений, коммуникаций, управления, данных, моделей, инструкций, знаний, значений, интеллектуальных воздействий, структур, восприятия и представления.

Существует две концепции понимания информации.

Физическая концепция. Информация – это фундаментальная категория, т.е. такая же основа мироздания, как и вещество, энергия. Известно выражение Н. Виннера, основателя теории информации: "Информация есть информация, а не вещество или энергия". Таким образом, Мир состоит из вещества, энергии и информации.

Коммуникативная концепция. Информация – все то, что может быть сообщено. Коммуникативной концепции различают два вида информации:

- существующая информация – сведения, которые можно сообщить о каком-либо объекте, явлении,
- передаваемая информация – сообщение по каналу информации.

Данные

Термин "Данные" происходит от латинского Datum – факт (англ. Data – данные). *Данные* – это совокупность фактов, представленных в формализованном виде (в количественном или качественном выражении).

Данные соответствуют дискретным зарегистрированным фактам относительно явлений. Данные относятся к информации или фактам обычно собранным как результат опыта, наблюдений или эксперимента или процессов в компьютерных системах или предположения. Данные могут состоять из чисел, слов, или изображений, особенно как результаты измерений или наблюдений

множества переменных. Данные часто представляют как нижний уровень абстракции, от которого получают информацию и знания. В результате обработки данных мы получаем информацию о реальном мире.

Понятие "данные" имеет ряд аспектов [27]:

- Данные могут быть первичными, объективно отвечающими в реальном мире предметам или условиям вне зависимости от способа фиксации и интерпретации человеком. Например, одно дерево выше другого, или река течет в направлении восхода Солнца независимо от того, наблюдает ли их человек.
- Данные могут быть зафиксированы тем или иным способом, не включающим интерпретацию, например, записаны на магнитном носителе электронным датчиком, как сейсмические сигналы, или на фотопленку, как фотографии ландшафта.
- Данные могут быть накоплены определенным образом, например, в виде бумажной карты с определенной легендой, в виде набора чисел - координат точек измерения глубин водного объекта и самих измеренных глубин.
- Данные могут быть интерпретированы человеком непосредственно в процессе их фиксации, например, полевая зарисовка рельефа, сделанная геодезистом, словесное описание обнажения горных пород в полевом дневнике геолога.
- Данные могут быть структурированы и/или организованы определенным образом, например, в виде таблиц в результатах переписи населения, или в виде пространственной базы данных под управлением географической информационной системы.

Соотнесение понятий информация, данные и знания

Данные мы можем полезно использовать, и тем самым увеличить наше знание о каком-либо явлении или предмете реального мира [27]. Если мы имеем в нашей базе данных множество каких-то чисел и не вооружены знанием того, что это за числа, к чему относятся, какие величины представляют и для чего предназначены, то эти числа не являются для нас полезной информацией. Тем не менее, эти непонятные цифры, хранимые в базе данных, тоже являются данными. Информация извлекается из них только с приложением нашего мышления или интуиции, базирующихся на некоторых знаниях о предмете. Например, в базе данных имеются некоторые пары чисел. Только знание того, что эти числа являются картографическими координатами, понимание того, что такое вообще система координат,

что такое картографические проекции и о какой картографической проекции идет речь в данном случае, как организована система координат в данной проекции - как направлены ее оси, где расположено начало координат, и, наконец, каковы масштабы по осям и в каких единицах представлены данные нам значения координат, позволяют придать смысл этим числам. Только зная все перечисленное, мы можем сказать, что данные, представленные этими числами, являются информацией о географическом положении каких-то объектов.

Таким образом, чтобы извлечь полезную информацию из отдельно взятых данных, не находящихся в каком-либо смысловом контексте, необходима некоторая дополнительная информация особого, структурного плана, помогающая интерпретировать новые данные и ввести их в общую систему знаний об окружающем мире.

Иначе говоря, данные, пока мы не имеем информации о том, как они связаны с системой знаний, в эту систему не входят и не являются полезной информацией. Как только мы привлекаем сведения о структуре их связи с существующей системой знаний, они оказываются связанными с этой системой, становятся ее частью и тем самым увеличивают объем этой системы знаний, становятся полезной информацией. Таким образом, знания предполагают системность, структурированность информации, обеспечивая возможность описания не только свойств, но и взаимосвязей между явлениями и объектами реального мира. Ценность данных возрастает на пути от разобщенных элементов данных к полезной информации и к систематическим знаниям.

2.2.2 Географическая информация и географические данные

Географическая информация (Geographic Information) – информация об объектах и явлениях, которые непосредственно или опосредованно связаны с определенным местоположением относительно Земли.

Географические данные (Geographic Data) представляют собой единство геопространственных, семантических и временных данных географических объектов (Рис.2.2.1).

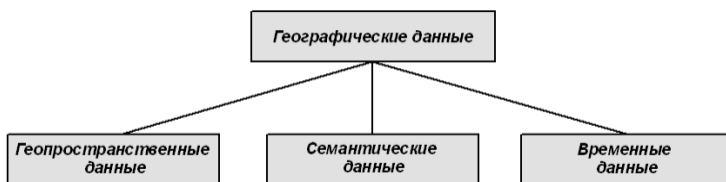


Рис. 2.2.1. - Состав географических данных

Геопространственные данные – это данные о локальных пространственных свойствах: местоположении, форме, размерах, и пространственных отношениях географических объектов, явлений, процессов в реальном земном пространстве. 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 их местоположение (позиционирует) на земной поверхности или в некой заданной системе координат. Основное требование к пространственному положению – точность.

Семантические данные – это данные, которыми описывается содержательная, смысловая информация о географических объектах, свойства географических объектов. Свойство – это категория, выражающая такую сторону объекта (сущности), которая обуславливает его различие или общность с другими объектами и обнаруживает себя при сопоставлении различных объектов. Свойства типа объектов выделяют из множества всех типов объектов подмножество объектов конкретного типа и являются постоянными для всех объектов этого типа. Свойства разных объектов конкретного типа различаются своими значениями: так, например, здания могут иметь различное число этажей, материал постройки, назначение и т. п. Тематический аспект (тема) обусловлен наличием признаков определенной тематики или предметной области. Это могут быть экономические, статистические, технические, организационные, управленческие и прочие виды данных. Требование к тематическим данным "полнота" означает, что этих данных достаточно для решения практических задач и нет необходимости проводить дополнительный сбор данных.

Временные данные фиксируют время исследования объекта и показывают изменение свойств объекта с течением времени. Основное

требование к временным данным - актуальность. Это означает, что актуальные данные можно использовать для обработки. Неактуальные данные - это устаревшие данные, которые нельзя полностью применять в новых изменившихся условиях.

В большинстве геоинформационных технологий для представления параметров времени и тематической направленности используют один класс данных - атрибуты.

Географическая информационная система должна быть способна управлять совместно частями географических данных.

2.2.3 Атрибуты пространственных объектов

Свойства географических объектов представляются в базах данных набором атрибутов. Атрибут (attribute) - синоним реквизита - свойство, качественный или количественный признак, характеризующий пространственный объект, и ассоциированный с его уникальным номером или идентификатором. Наборы значений атрибутов (attribute value) обычно представляются в форме таблиц реляционных баз данных. При этом строка (запись) представляет атрибуты одного объекта, а столбец (поле) - атрибуты одного типа. Для упорядочения, хранения и манипулирования атрибутивными данными используются средства систем управления базами данных СУБД.

Для атрибутов важнейшей характеристикой является тип использованной шкалы измерений [11], [13]. Атрибуты могут быть категоризованы по шкалам измерения данных. Шкалы измерения данных разделяются на шкалы отношений, интервальные, порядковые и номинальные. Общепринято деление шкал (и, соответственно, данных в этих шкалах) на качественные и количественные. К количественным относят интервальную (интервалов) и рациональную (отношений) шкалы. К качественным относят шкалы номинальную (наименований) и ординальную (порядковую, ранговую).

Данные отношений - данные, которые представляют количества в условиях равных интервалов и точки абсолютного нуля от начала. Представляются числами, которые могут быть преобразованы или объединены с любой математической функцией, чтобы генерировать значащие результаты. Примеры: измерения характеристик типа возраста, частоты, физических расстояний и денежных значений.

Интервальные данные - данные, которые представляют количества в условиях равных интервалов или степеней разности, но

чья нулевая точка (или точка начала) произвольно установлена. Интервальные значения данных и соотношение значений данных позволяют делать вычисления. Используются для определения количества различия, но не пропорций и характеристик относительно положения в пространстве, времени или размера. Сравнения могут делаться с более точной оценкой различий, чем в случае порядковых данных. Примеры: широты, долготы, направления, времена дня, нормализованное множество.

Порядковые данные - данные, которые определяют количество различий порядком, а не величиной. Размер интервалов не определен. Используются там, где количественные различия очевидны, когда величины представлены упорядоченным порядком. Примеры: различие типа "больше" или "меньше", различие между "бедными", "умеренными" и "хорошими" сельскохозяйственными землями (не дают никаких указаний, насколько точно лучший / худший каждый тип, чем другой).

Номинальные данные - данные, которые обеспечивают имя или идентификатор, представляют качества, а не количества. Объекты различаются по именам. Система позволяет делать высказывания о том, как называется объект, но не позволяет делать прямого сравнения одного объекта и другого, за исключением определения тождества. Примеры: телефонные номера, почтовые коды или типы деревьев, типа покрытия земли (1 = пашня, 2 = лес, 3 = болото, 4 = застроенная территория) это просто удобная форма именования (номинального значения).

Кроме этого, атрибуты могут быть первичными (измеренными, введенными) и вторичными, расчетными, полученными расчетом из значений других атрибутов. Частный случай таких расчетных - это атрибуты (обычно пространственные), которые рассчитываются исходя из позиционных параметров объектов (например, периметр). Некоторое свойство, непрерывно распределенное в пространстве, например, на земной поверхности, удобно в математическом смысле рассматривать в качестве поля (вообще говоря, разного вида - скалярного, векторного, тензорного, трехмерного или двумерного, определенного только на поверхности земли или с ней не связанного).

2.2.4 Виды компьютерных моделей географических объектов

В ГИС географические объекты могут быть представлены следующими распространенными компьютерными моделями:

Векторные модели – это модели географических объектов на основе представления векторов совокупностями координатных пар.

Растровые модели - это модели географических объектов в виде совокупности ячеек регулярной сетки или растра.

Триангуляционные модели - это модели географических объектов в виде сети смежных треугольников.

2.2.5 Контрольные вопросы и задания для самостоятельной работы

- 1) Какое соотношение понятий информация, данные и знания?
- 2) Что есть Географическая информация и Географические данные?
- 3) Приведите характеристику компонентов географических данных.
- 4) Что понимается под атрибутами пространственных объектов?
- 5) Дайте характеристику различных шкал измерения атрибутов.

Раздел 2.3

ВЕКТОРНЫЕ МОДЕЛИ ГЕОГРАФИЧЕСКИХ ОБЪЕКТОВ

2.3.1 Базовые понятия картографического представления географических объектов

Карта – графическое представление географических объектов. Карта – наиболее общий метод, используемый для представления пространственной информации. Метод, с которым все знакомы. Однако несколько ключевых понятий картографического отображения составляют основу понимания того, как в ГИС представляется географическая информация.

Пространственная информация о географических объектах представляется на карте наборами элементарных графических примитивов – точек, линий, областей (полигонов).

- Точками представляют такие географические объекты, размерами которых для конкретной цели можно пренебречь. Например, колодец подземной коммуникации, опора линии электропередачи, человек для демографического изучения ситуации, остановки городского пассажирского транспорта для изучения транспортного обслуживания...
- Линиями представляют такие географические объекты, которые являются настолько узкими, что имеют длину, но не имеют площади. Например, дороги, водные потоки, инженерные коммуникации, границы, ...
- Областями (замкнутыми внутри полигона фигурами) представляют такие географические объекты, которые имеют форму, местоположение и площадь. Например, здание, квартал, проезжая часть улицы, ...

Семантическая информация на картах передается условными знаками (символами и цветом) и надписями (например, название улицы, этажность дома...). Несмотря на большое разнообразие используемых условных знаков и надписей, насыщение карт семантической информацией ограничено из-за отсутствия свободного места на карте для ее отображения. Чрезвычайная насыщенность условных знаков и надписей приводит к потере возможности восприятия информации с карты.

2.3.2 Простые векторные модели географических объектов

Вектор определяется как отрезок прямой, которому присвоено определенное положительное направление. Вектор характеризуют начало, конец, длина, направление. Вектор может быть свободным и несвободным. Операции над векторами: сложение, вычитание, умножение скалярное, умножение векторное, умножение смешанное. Векторный аппарат положен в основу создания векторных моделей географических объектов.

2.3.2.1 Простые векторные модели единичных географических объектов

В ГИС представление географических объектов [28] выполняется посредством отображения их геометрической формы на двумерной плоскости с использованием элементарных графических примитивов:

- точек (нуль-мерных векторов),
- линий (полилиний, образованных плоскими двумерными векторами),
- областей (форм, ограниченных полигонами – замкнутой последовательностью двумерных векторов); географические объекты, представляемые областями, называют полигональными объектами.

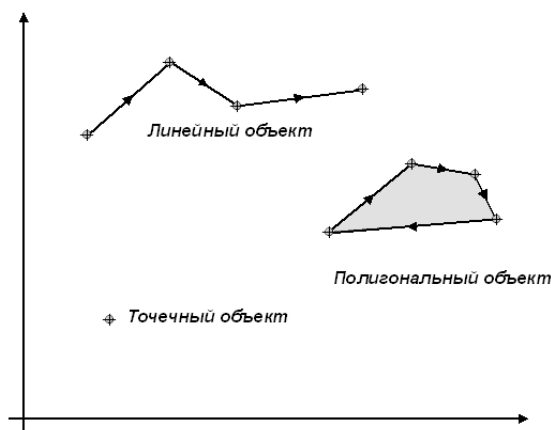


Рис. 2.3.1 - Векторные модели представления единичных географических объектов

Каждое векторное образование в цифровой форме представляется координатными парами X, Y .

- Точечный географический объект представляется одной координатной парой X, Y .
- Линейный географический объект представляется последовательностью координатных пар $X_1, Y_1; X_2, Y_2; X_3, Y_3; \dots$ сегментов полилинии.
- Полигональный географический объект представляется последовательностью координатных пар $X_1, Y_1; X_2, Y_2; X_3, Y_3; \dots; X_1, Y_1$ сегментов полигона. В этом списке совпадение первой и последней пары координат означает, что полигон замкнутый.

2.3.2.2 Простые векторные модели множества географических объектов

При наличии географических объектов числом 2 и более каждому присваивается номер.

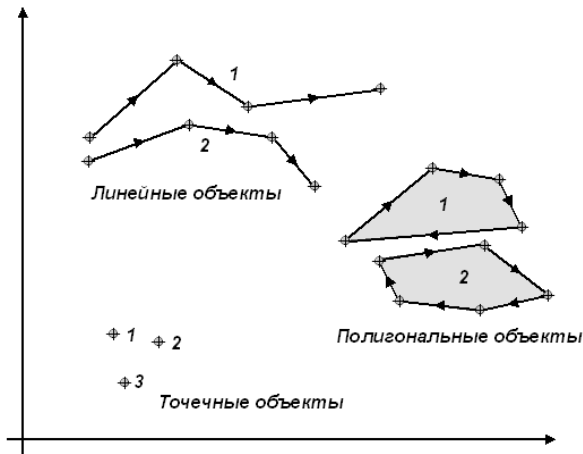


Рис. 2.3.2 - Векторные модели представления множества географических объектов

В этом случае последовательностям координатных пар присваивается номер географического объекта. Каждое векторное образование в цифровой форме представляется координатными парами X, Y .

Точечные географические объекты представляются таблицей координатных пар (табл.2.3.1).

Таблица 2.3.1 - Таблица точечных географических объектов

Номер географического объекта	Координатные пары
1	X_1, Y_1
2	X_2, Y_2
3	X_3, Y_3
4	X_4, Y_4
5	X_5, Y_5

Линейные географические объекты представляются таблицей последовательностей координатных пар (табл.2.3.2).

Таблица 2.3.2 - Таблица линейных географических объектов

Номер географического объекта	Координатные пары
1	$X_{11}, Y_{11}, X_{12}, Y_{12}, X_{13}, Y_{13}$
2	$X_{21}, Y_{21}, X_{22}, Y_{22}, X_{23}, Y_{23}, X_{24}, Y_{24}, X_{25}, Y_{25}$
3	$X_{31}, Y_{31}, X_{32}, Y_{32}, X_{33}, Y_{33}, X_{34}, Y_{34}$

Полигональные географические объекты представляются таблицей последовательностей координатных пар (табл.2.3.3)

Таблица 2.3.3 - Таблица полигональных географических объектов

Номер географического объекта	Координатные пары
1	$X_{11}, Y_{11}, X_{12}, Y_{12}, X_{13}, Y_{13}, X_{11}, Y_{11}$
2	$X_{21}, Y_{21}, X_{22}, Y_{22}, X_{23}, Y_{23}, X_{24}, Y_{24}, X_{25}, Y_{25}, X_{21}, Y_{21}$
3	$X_{31}, Y_{31}, X_{32}, Y_{32}, X_{33}, Y_{33}, X_{34}, Y_{34}, X_{31}, Y_{31}$
4	$X_{41}, Y_{41}, X_{42}, Y_{42}, X_{43}, Y_{43}, X_{41}, Y_{41}$

В ГИС для формирования векторных моделей используются такие координаты, которые представляют местоположение географических объектов в реальном земном пространстве: географические координаты на сфероиде (широта, долгота) или Декартовы координаты на плоскости (абсцисса, ордината), полученные в определенной картографической проекции.

Концептуально здесь то, что географические объекты в компьютере хранятся как файлы координатных пар, как наборы цифр, т.е. в цифровой форме. Табличные списки координатных пар есть форма представления плоских цифровых файлов.

Создание векторных файлов возможно 1) путем ввода координат клавиатурой, 2) в результате процесса оцифровки или "дигитализации" (англ. *Digit* - цифра), т.е. трансформирования аналоговой информации графических объектов карты в цифровой вид.

2.3.3 Топологические векторные модели географических объектов

Для изучения этого вопроса необходимо предварительно рассмотреть базовые понятия "Граф", "Пространственные отношения", "Топология".

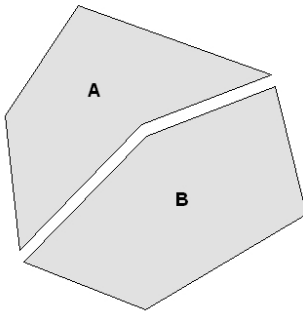
2.3.3.1 Понятие "Граф"

Изложенное представление географических объектов простыми векторными моделями в виде списков координатных пар позволяет образовывать сложные объекты в виде комбинации элементарных объектов – точек, линий, полигонов.

Однако представление смежных полигонов простыми векторными моделями становится не эффективным потому, что стороны, общие для двух смежных полигонов, хранятся дважды. Более эффективный способ хранения данных основан на понятии "Граф".

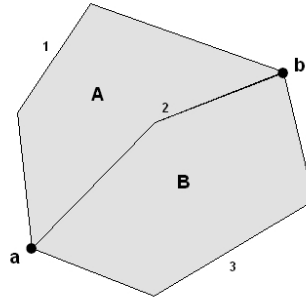
Граф - есть множество элементов, связанных между собой отношениями. Геометрически граф представляется в форме векторной схемы, состоящей из вершин, узлов, ребер, дуг.

- *Вершина* (англ. Vertex) – это объект графа. Вершины представляют точками.
- *Ребро* (англ. Edge) – это линия, которая связывает точки (объекты графа). Ребра представляют отношения между объектами.
- *Дуга* (англ. Arc) – это ребро с определенной ориентацией относительно ее конечных вершин.
- *Узел* (англ. Node) – это вершина, общая для двух и большего числа дуг. В узлах сходятся дуги.



a)

Рис. 2.3.3 а) - Простое векторное представление смежных полигонов.



b)

Рис. 2.3.3 б) - Отображение полигонов на основе структуры "граф"

На рисунках изображены два альтернативных метода для представления полигонов.

- Первый метод (рис. 2.3.3 а) описывает полигон как замкнутую последовательность векторов, представленную списком координат.
- Второй метод (рис. 2.3.3 б) описывает полигон как набор дуг (1,2,3) и узлов (a,b). Полигон А определяют дуги 1,2. Полигон В определяют дуги 2,3.

В первом случае координаты точек смежных линий повторяются. Во втором случае общую границу смежных полигонов представляет дуга 2, координаты которой используются один раз при описании дуги.

2.3.3.2 Понятие "Пространственные отношения"

Пространственные отношения – это одно из фундаментальных свойств пространства. *Пространственные отношения* – это такие отношения, с помощью которых описывают пространственные взаимосвязи объектов. Существует большое количество видов пространственных отношений. Примеры видов пространственных отношений: находиться близко, далеко, соприкасаться, находиться

внутри, вне, пересекаться, примыкать, находиться выше, ниже, левее, правее...

На определенной территории существует огромное множество пространственных отношений. О количестве пространственных отношений на карте говорит следующий факт. В 1 кв. см. карты масштаба 1:10000 при дискретизации 0,1 мм. содержится порядка 10^{70} – 10^{80} различных пространственных отношений (порядка, эквивалентности, различия, сходства...). Для сравнения, масса планеты Земля составляет 10^{23} грамм.

На картах пространственные отношения представлены неявно, скрыто. Пространственные отношения воспринимаются, оцениваются, интерпретируются тем, кто читает карту. Объем восприятия информации зависит от подготовки, квалификации субъекта. Можно оценить, например, какие объекты находятся вблизи дороги, пересекаются ли две магистрали, какие наиболее высокие места, какой кратчайший маршрут, и т.д.

Пространственные отношения воспринимаются как непосредственно, так и опосредованно. В любом случае огромный объем информации, содержащийся на картах, остается невостребованным.

2.3.3.3 Понятие "Топология"

Топология – это раздел математики, изучающий идею непрерывности. Непрерывность – это одно из фундаментальных свойств категории пространства-времени. Топология изучает пространственные отношения, которые не изменяются при любых непрерывных преобразованиях пространства. Топология реализуется математической процедурой явного определения пространственных отношений.

В ГИС топология реализуется на основе оригинальной и простой идеи: различные типы пространственных отношений представляются связанными списками объектов. В ГИС реализованы три базовых топологических отношения дуг (*Arc*):

1. Дуги, которые соединяются в полигон, окружают область;
2. Дуги, имеющие направление, имеют правые и левые стороны.
3. Дуги соединяются в узлах;

На этой основе в ГИС [28] формализована топологическая группа пространственных отношений, которые не изменяются при любых непрерывных преобразованиях пространства:

1. Область (*Area*) – определяется дугами, которые соединяются в полигон для окружения области.
2. Смежность (*Contiguity*) – определяется дугами, имеющими направление, и поэтому имеют правую и левую стороны.
3. Связность (*Connectivity*) – определяется дугами, которые соединяются в узлах.

Создание и хранение моделей географических объектов на основе пространственных отношений имеет ряд преимуществ:

- Данные хранятся более эффективно,
- Данные можно обрабатывать быстрее и большими наборами,
- Топология облегчает аналитические функции.

2.3.3.4 Топологическое представление области

Топологическое представление области рассмотрим на примере карты областей (рис. 2.3.4)

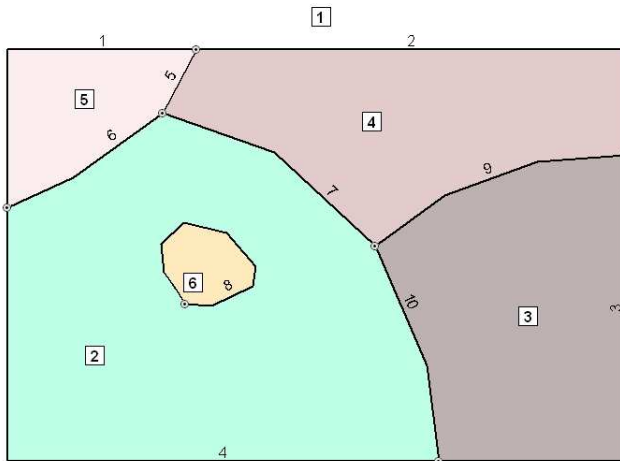


Рис. 2.3.4 Карта областей – полигональных объектов

На карте:

- полигон 1 – внешний полигон
- полигон 2 охватывают дуги 4,6,7,10,8;
- полигон 3 охватывают дуги 3,9,10;
- полигон 4 охватывают дуги 2,7,9;
- полигон 5 охватывают дуги 1,5,6;
- полигон 6 охватывает дуга 8.

В ГИС топологическое представление области реализуется списком "Полигон – Дуга" и связанным с ним списком "Координаты дуг".

Табл. 2.3.4 - Список "Полигон – Дуга"

Полигон	Дуга
1	1,2,3,4,0
2	4,6,7,10,0,8
3	3,9,10
4	2,7,9
5	1,5,6
6	8

Табл. 2.3.5 - Список "Координаты дуг"

Дуга	Координаты вершин
1	$X_{11}, Y_{11}, X_{12}, Y_{12}, X_{13}, Y_{13}$
2	$X_{13}, Y_{13}, X_{21}, Y_{21}, X_{22}, Y_{22}$
3	$X_{22}, Y_{22}, X_{31}, Y_{31}, X_{32}, Y_{32}$
4	$X_{33}, Y_{33}, X_{41}, Y_{41}, X_{11}, Y_{11}$
5	$X_{15}, Y_{15}, \dots, X_{13}, Y_{13}$
6	$X_{15}, Y_{15}, \dots, X_{11}, Y_{11}$
7	$X_{71}, Y_{71}, \dots, X_{15}, Y_{15}$
8	$X_{81}, Y_{81}, \dots, X_{88}, Y_{88}$
9	$X_{71}, Y_{71}, \dots, X_{22}, Y_{22}$
10	$X_{71}, Y_{71}, \dots, X_{32}, Y_{32}$

В списке "Полигон – Дуга" полигон 2 описан дугами 4,6,7,10,0,8, где 0 перед дугой 8 указывает, что эта дуга создает остров (дыру) в полигоне 2. В списке "Полигон – Дуга" дуги могут появляться 2 раза, в списке координат дуг – один раз.

Конечными вершинами дуг являются узлы. Узел двух и большего числа дуг должен иметь в списке одни координаты.

Таким образом, топологическое описание области реализуется в цифровой форме двумя связанными списками.

2.3.3.5 Топологическое представление смежности

Топологическое представление смежности рассмотрим на примере карты дуг (Рис. 2.3.5)

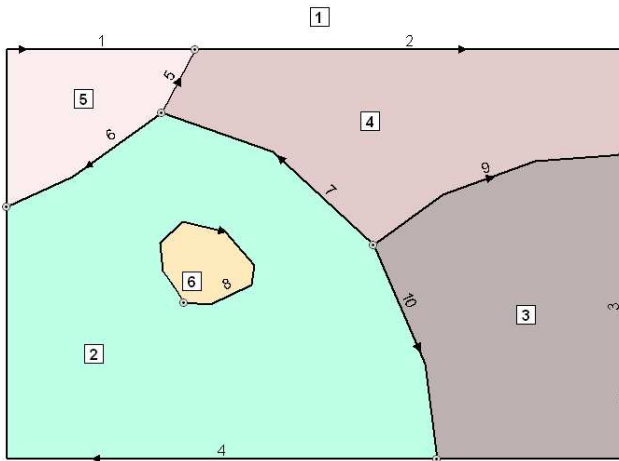


Рис. 2.3.5 - Карта дуг и полигонов

На карте

- слева относительно дуги 5 находится полигон 5, а справа – полигон 4;
- слева относительно дуги 6 находится полигон 2, а справа – полигон 5;
- слева относительно дуги 1 находится полигон 1, а справа – полигон 5;
- и т.д.

Снаружи всех полигонов находится внешний полигон 1, называемый полигоном Вселенной. Он вводится для однообразного описания полигонов: каждая дуга должна иметь полигон слева и полигон справа.

Топологическое отношение смежности позволяет идентифицировать кто рядом. Например: кто соседи земельного участка? Лес смежный с озером?

Любые полигоны, совместно использующие общую дугу, являются смежными. Так как дуга имеет направление от узла к узлу, можно поддерживать список полигонов с левой и с правой стороны.

В ГИС топологическое представление смежности реализуется списком "Полигон слева-справа" и связанным с ним списком "Координаты дуг".

Табл. 2.3.6 - Список
"Полигон слева-справа"

Дуга	Полигон слева	Полигон справа
1	1	5
2	1	4
3	1	3
4	1	2
5	5	4
6	2	5
7	2	4
8	2	6
9	4	3
10	3	2

Табл. 2.3.7 - Список
"Координаты дуг"

Дуга	Координаты вершин
1	$X_{11}, Y_{11}, X_{12}, Y_{12},$ X_{13}, Y_{13}
2	$X_{13}, Y_{13}, X_{21}, Y_{21},$ X_{22}, Y_{22}
3	$X_{22}, Y_{22}, X_{31}, Y_{31}, X_{32},$ Y_{32}
4	$X_{33}, Y_{33}, X_{41}, Y_{41}, X_{11},$ Y_{11}
5	$X_{15}, Y_{15}, \dots, X_{13}, Y_{13}$
6	$X_{15}, Y_{15}, \dots, X_{11}, Y_{11}$
7	$X_{71}, Y_{71}, \dots, X_{15}, Y_{15}$
8	$X_{81}, Y_{81}, \dots, X_{88}, Y_{88}$
9	$X_{71}, Y_{71}, \dots, X_{22}, Y_{22}$
10	$X_{71}, Y_{71}, \dots, X_{32}, Y_{32}$

Таким образом, топологическое описание смежности реализуется двумя списками в цифровой форме.

2.3.3.6 Топологическое представление связности

Топологическое представление связности рассмотрим на примере карты дуг (рис. 2.3.6)

Конечные точки дуги называются "узлами". Каждая дуга имеет два узла: начальный, называемый "из-узла", и конечный, называемый "к-узлу". Дуги могут соединяться только в узлах.

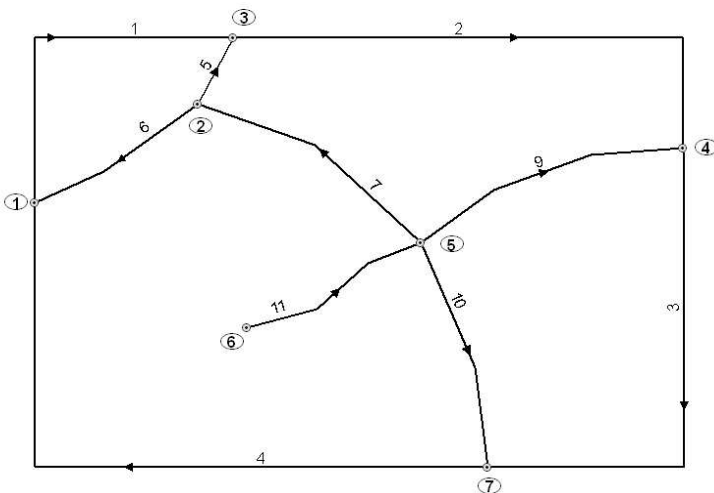


Рис. 2.3.6 - Карта дуг и узлов

На рис. 2.3.6:

- дуги 7,8,9,10 соединяются в узле 5;
- дуги 5,6,7 соединяются в узле 2;
- и т.д.

В ГИС топологическое представление связности реализуется списком "Дуга-Узел" и связанным с ним списком координат дуг.

Табл. 2.3.8 - Список
"Дуга-Узел"

Дуга	От-узла	К-узлу
1	1	3
2	3	4
3	4	7
4	7	1
5	2	3
6	2	1
7	5	2
8	8	8
9	5	4
10	5	7
11	6	5

Табл. 2.3.9 - Список
"Координаты дуг"

Дуга	Координаты вершин
1	$X_{11}, Y_{11}, X_{12}, Y_{12}, X_{13}, Y_{13}$
2	$X_{13}, Y_{13}, X_{21}, Y_{21}, X_{22}, Y_{22}$
3	$X_{22}, Y_{22}, X_{31}, Y_{31}, X_{32}, Y_{32}$
4	$X_{33}, Y_{33}, X_{41}, Y_{41}, X_{11}, Y_{11}$
5	$X_{15}, Y_{15}, \dots, X_{13}, Y_{13}$
6	$X_{15}, Y_{15}, \dots, X_{11}, Y_{11}$
7	$X_{71}, Y_{71}, \dots, X_{15}, Y_{15}$
8	$X_{81}, Y_{81}, \dots, X_{88}, Y_{88}$
9	$X_{71}, Y_{71}, \dots, X_{22}, Y_{22}$
10	$X_{71}, Y_{71}, \dots, X_{32}, Y_{32}$
11	$X_{61}, Y_{61}, \dots, X_{71}, Y_{71}$

Таким образом, топологическое описание связности в цифровой форме реализуется 2 списками.

При прослеживании всех дуг в списке "Дуга-Узел" программа определяет, какие дуги соединяются (связаны) друг с другом.

Связность эффективна для решения транспортных задач. Например, можно проехать по дугам 6,7,10 через узлы 5,2, но нельзя переехать непосредственно с дуги 6 на дугу 10, которые не имеют общего узла.

2.3.3.7 Сравнение векторных моделей географических объектов

Существует два основных типа векторных моделей географических объектов: простые нетопологические модели (называемую моделью "spaghetti" из-за неупорядоченных векторных объектов) и топологические модели.

Табл. 2.3.10 - Сравнение векторных моделей географических объектов

	Простые нетопологические модели (модель "spaghetti")	Топологические модели
1	Возможны незамкнутые полигоны	Однозначное представление области
2	Возможно отсутствие связности линий	Дуги связаны через узлы
3	Неэффективное хранение данных	Эффективное хранение данных
4	Отсутствует возможность анализа данных	Возможность анализа данных

Таким образом, топологические модели обладают существенными преимуществами по сравнению с моделями "spaghetti".

2.3.4 Форматы векторных данных

Векторные графические форматы данных представлены в табл. 2.3.11.

Таблица 2.3.11 - Векторные графические форматы данных

Наименование формата	Описание
DXF, DWG, DGN	Форматы данных систем автоматизированного проектирования (САПР)
DX90	Формат цифровых навигационных карт
DLG	Формат данных геологической съемки США
DWF	Формат передачи графических данных по Интернету
FIM	Формат данных Роскартографии, предназначен для обмена данными
GEN	Обменный формат ГИС-пакета ARC/INFO
SHP	Формат данных ГИС-пакета Arc View (шейп-файл), описывается несколькими файлами записей с определенными расширениями: .shp — позиционные данные; .shx — индекс формы пространственных данных; .dbf — атрибутивные данные и др.
TAB	Формат ГИС-пакета MapInfo; описывается файлами с определенными расширениями: .tab — текстовое описание структуры данных таблиц; .dat — табличные данные; .map — графические объекты; .ID — список указателей (индекс) на графические объекты
MIF/MID	Обменный формат ГИС-пакета MapInfo
HPGL	Формат вывода на принтер или графопостроитель
DMF	Формат ГИС-пакета Digital

Среди векторных наибольшее распространение в ГИС получил формат DXF пакета AutoCad (Autodesk Inc.), использующий для передачи графической атрибутивной информации формат DBF (dBase).

2.3.5 Резюме представления географических объектов векторными моделями

Обобщенная характеристика представления географических объектов векторными моделями приведена в табл. 2.3.12.

Таблица 2.3.12 – Резюме представления географических объектов векторными моделями [29]

Фокус модели	Векторные данные ориентированы на моделирование дискретных пространственных объектов с точным определением форм и границ
Источники данных	Компиляция материалов аэрофотосъемки, сбор GPS определений, оцифровка бумажных карт, оцифровка по дисплею, векторизация растровых данных, построение изолиний рельефа, обработка данных топографических съемок, импорт из CAD чертежей
Хранение пространственных данных	Точки хранятся как координатные пары; Линии - как последовательность координатных пар, Полигоны – как замкнутая последовательность координатных пар
Представление пространственных объектов	Точками представляют такие географические объекты, размерами которых для конкретной цели можно пренебречь. Линиями представляют такие географические объекты, которые являются настолько узкими, что имеют длину, но не имеют ширину. Области представляют такие географические объекты, которые имеют местоположение, форму и площадь.
Топологические отношения	Топология линий опирается на принцип, по которому линии связываются в узлах. Топология полигонов опирается на принцип, по которому полигоны находятся слева и справа от линии.

Продолжение таблицы 2.3.12

Геопространственный анализ	Топологический оверлей карт, генерирование буферов и анализ близости, растворение полигонов и оверлей, пространственные и логические запросы, адресное геокодирование, сетевой анализ.
Картографическая продукция	Векторные данные являются лучшими для вычерчивания точной формы и положения пространственных объектов. Они не совсем пригодные для непрерывных явлений или пространственных объектов с нечеткими границами.

2.3.6 Контрольные вопросы и задания для самостоятельной работы

- 1) Какие ключевые понятия картографического отображения используются в ГИС?
- 2) В чем суть представления в ГИС географических объектов простыми векторными моделями?
- 3) Приведите определения базовых понятий "Граф", "Пространственные отношения", "Топология".
- 4) Как формируется векторное топологическое представление области?
- 5) Как формируется векторное топологическое представление смежности?
- 6) Как формируется векторное топологическое представление связности?
- 7) Сравните простые нетопологические и топологические векторные модели географических объектов.
- 8) Приведите краткую характеристику векторных графических форматов данных.
- 9) Приведите обобщенную характеристику представления географических объектов векторными моделями.

Раздел 2.4

РАСТРОВЫЕ МОДЕЛИ ГЕОГРАФИЧЕСКИХ ОБЪЕКТОВ

2.4.1 Концепция растровых моделей географических объектов

2.4.1.1 Определение растровых моделей

Земное пространство непрерывно. Для представления земного пространства используются модели данных, которые в любом случае основаны на наборах дискретных объектов.

В *векторной* модели дискретизация реального земного пространства выполняется посредством выделения дискретных географических объектов и отображения их отдельными пространственными элементами – точками, линиями, областями.

В *растровой* модели дискретизация реального земного пространства достигается путем разделения его непрерывной последовательности на множество смежных элементарных объектов – пространственных ячеек. Ячейки являются базовыми элементами растровой модели данных.

Ячейка – это наименьшая единица информации растровой модели. Под ячейкой понимается гомогенный объект, т.е. такой объект, который имеет одну характеристику или одно значение. Поэтому растровая модель дает информацию о том, что расположено в данном месте. Для сравнения векторная модель дает информацию о том, где расположен объект. Для каждой растровой модели используются ячейки одного выбранного размера. В ГИС в зависимости от способа задания ячеек различают два типа пространственных ячеек:

- *пиксел* (Pixel - Picture Element) – ячейка, представляющая собой минимальный неделимый далее элемент изображения;
- *ячейка* (Cell) – ячейка заданного размера в форме квадрата, прямоугольника, треугольника, шестиугольника или трапеции. Наиболее распространенная форма ячейки - квадрат или прямоугольник.

Для отображения непрерывных последовательностей реального мира используются растровые модели данных.

Растр (Raster, Tessellation) в ГИС – модель географического пространства в виде регулярной матрицы смежных ячеек, локализованных в реальном земном пространстве.

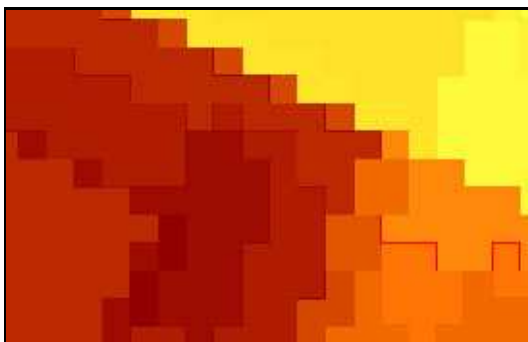


Рис. 2.4.1 – Растр

Для хранения растровых данных о географическом пространстве используют разные форматы данных. Одним из них является формат данных ESRI, называемый *grid* (ESRI Grid). Растр, образованный пикселями, представляет изображение (Image).

2.4.1.2. Источники растровых данных

Источниками растровых данных являются :

- фотоизображения (Image):
 - аэрофотоснимки территории;
 - космические снимки территории;
 - фотографии объектов;
- чертежи:
 - топографические карты;
 - планы;
 - технические чертежи;
 - схемы;
- рисунки;
- тексты:
 - документы;
 - таблицы.

Источники растровых данных преобразовываются в цифровой вид путем их сканирования.

2.4.1.3 Матрица ячеек

Любая растровая модель в общем случае представляется регулярной матрицей ячеек.

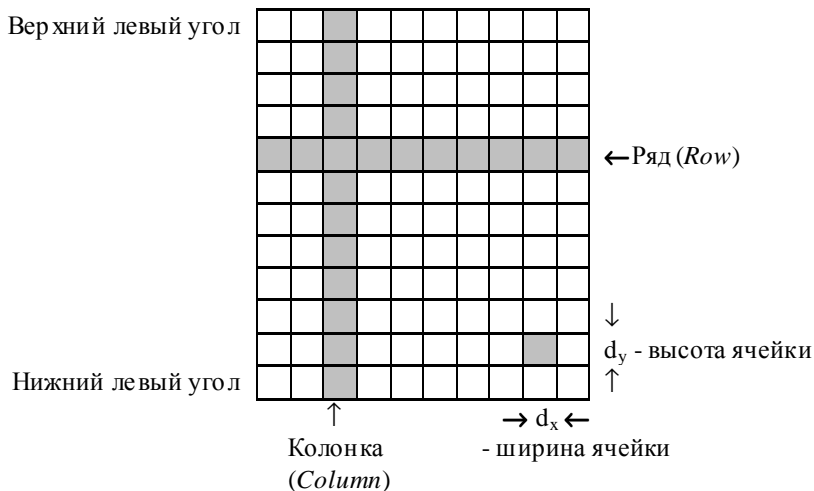


Рис. 2.4.2 – Матрица ячеек

Матрица ячеек имеет ряды ячеек и колонки ячеек. Положение каждой ячейки обозначается номером C колонки (Column) и номером R ряда (Row) от начала отсчета. За начало отсчета принимают или верхний левый угол или нижний левый угол матрицы.

Каждая ячейка имеет ширину d_x (размер по оси X), и высоту d_y (размер по оси Y). Ячейки одного растра имеют одинаковые размеры.

2.4.1.4 Представление дискретных географических объектов

В растровых моделях для представления дискретных географических объектов используются площадные объекты – ячейки растра.

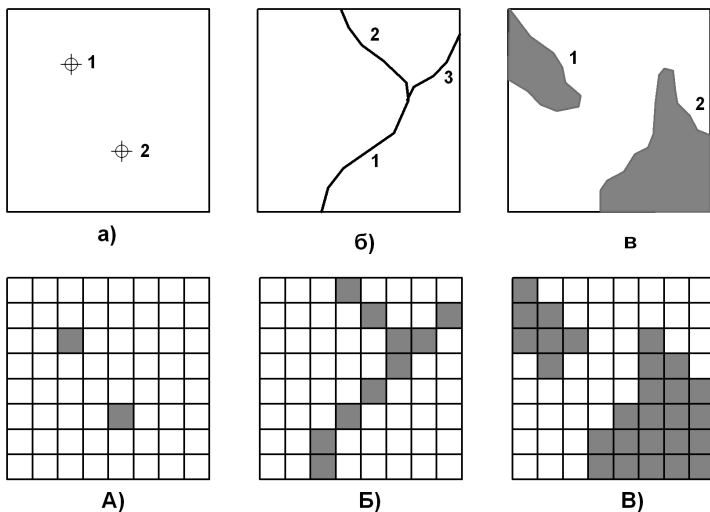


Рис. 2.4.3 - Представление географических объектов векторными моделями а), б), в) и растровыми моделями А) Б) В),

Географические объекты отображаются ступенчатым образом.

Точечные объекты представляются отдельными ячейками (рис. 2.4.3, А). Точечный 0-мерный объект, не имеющий размера, отображается двумерной структурой, имеющей ширину и длину.

Линейные объекты представляются цепочкой соединенных ячеек (рис. 2.4.3 , Б). Линейные объекты, являющиеся одномерными, отображаются двумерной структурой ячеек.

Области представляются набором смежных ячеек, который является двумерной структурой (рис. 2.4.3 , В). Границей области есть линейный объект – полигон.

Растровая модель реального земного непрерывного пространства представляет не только отдельные географические объекты, но и ту часть пространства, в которой их нет.

2.4.1.5 Представление непрерывных явлений регулярными моделями

Определение поверхности

Непрерывные явления (поля) отображаются поверхностями. Поверхность можно определить как кусок плоскости, подвергнутой непрерывным деформациям (растяжениям, сжатиям, изгибаниям), каждая точка которой представляется значениями координат X , Y , Z . Значения Z поверхности меняются при переходе от одного места к другому, образуя непрерывности пространственного градиента.

Если параметр Z ассоциируется с высотой, то определенная на нем поверхность будет представлять рельеф местности. В ГИС параметр Z могут представлять любые измеримые величины, такие, как стоимость, атмосферное давление, плотность населения, температура и так далее. Определенные на них поверхности будут соответственно поверхностями стоимости, давления, плотности населения, температуры, ... В тех случаях, когда значение Z можно трактовать как статистическое представление величины рассматриваемых явлений и объектов, образованную на таких величинах поверхность называют "статистической поверхностью".

В ГИС для моделирования поверхности используют регулярные модели (растровые и сеточные) и нерегулярные модели (триангуляционные модели).

Растровое представление поверхности

Растры представляют поверхности в виде регулярной матрицы двумерных ячеек со значением Z . Каждая ячейка хранит свое значение Z . Трехмерная растровая модель поверхности представляет собой совокупность смежных блоков. Значения Z отображаются соответствующим цветом (Рис. 2.6.4).

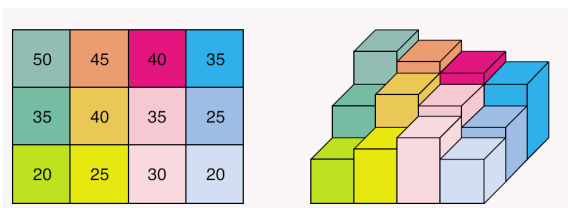


Рис. 2.4.4 – 2D и 3D растровое представление поверхности [29].

Преимущества растрового представления поверхности: 1) это простая модель; 2) применение в приложениях, где точность местоположений не имеет преимущественного значения и не требуется точного представления пространственных объектов поверхности.

Недостатки растрового представления поверхности: 1) регулярная структура не приспособлена к изменениям сложного рельефа; 2) разрывы непрерывности передаются недостаточно хорошо; 3) точные местоположения точек вершин, дна теряются.

Сеточное представление поверхности

С растровым представлением поверхности связано сеточное представление поверхности. Каждую ячейку растра идентифицируют местоположением и значением Z . Для определенных типов данных значение ячейки представляет собой измеренное значение в центральной точке ячейки.

Для сеточного представления поверхности базовой структурой является сетка, образованная параллельными и взаимно перпендикулярными линиями, проходящими через центральные точки ячеек растра. В ГИС для отображения поверхности на основании такой сетки используются две основные модели – грид и латтис [29], [30], [31], [32].

Грид (Grid) – модель в виде коллекции центральных точек ячеек со значениями Z , расположенных регулярно через горизонтальные интервалы (в отличие от формата ESRI Grid). Грид передает форму поверхности точками в трехмерном пространстве. Если значениями Z являются высоты местности, грид представляет собой цифровую модель высот (Digital elevation model – DEM).

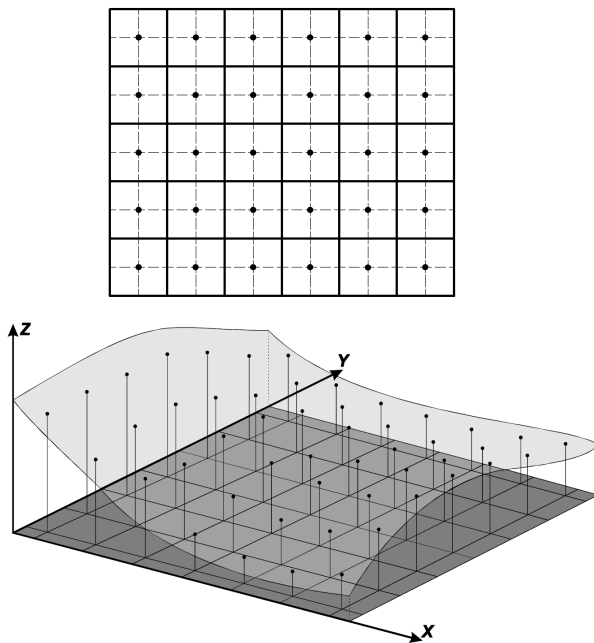


Рис. 2.4.5 – Грид в 2D и 3D представлении [33].

Латтис (Lattice) представляет собой структуру точек и линий пространственной решетки. Латтис передает форму поверхности, используя линии пространственной решетки. Если значениями Z являются высоты местности, латтис представляет собой геометрическую модель местности (Geometric terrain model - GTM).

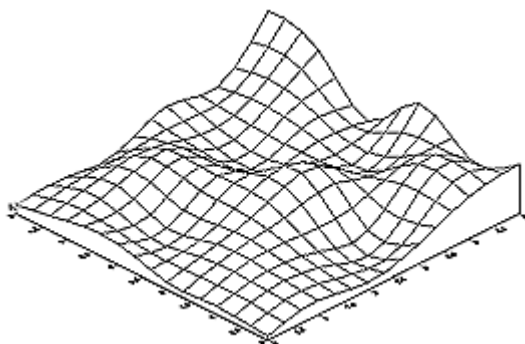


Рис.2.4.6 - Латтис в 3D представлении

В трехмерном представлении узел четырех линий передает значение Z. При этом четыре линии, связанные с этим узлом растягиваются пропорционально. Функции сглаживания линий позволяют представлять поверхность плавно изменяющейся формой (Рис.2.4.6). Применение латтис может обеспечить более правильную картину поверхности.

2.4.1.6 Недостатки и преимущества растровых моделей

По сравнению с векторными моделями растровые модели обладают следующими недостатками:

- географические объекты характеризуются менее точной информацией о местоположении и размерах;
- растры требуют больших объемов памяти.

Представление географических объектов растровыми моделями имеет следующие преимущества:

- растр отображает непрерывно охватываемую территорию;
- растровые данные проще для обработки и обеспечивают более высокое быстродействие;
- ввод растровых данных менее трудоемкий.

2.4.2 Характеристики растровых моделей

Для растровых моделей существует ряд характеристик: разрешение, значение, ориентация, положение, др.

2.4.2.1 Разрешение

Разрешение – размер наименьшего из различных участков пространства (поверхности), отображаемый одной ячейкой.

Более высоким разрешением обладает растр с меньшим размером ячеек. Высокое разрешение подразумевает обилие деталей, множество ячеек, минимальный размер ячеек. Чем больше размер ячейки, тем большую площадь она покрывает, тем меньшее разрешение растра.

Для разных задач используются разные размеры ячеек, растры разного разрешения. Например, для изучения использования земли (земли промышленности, земли многоэтажной, среднеэтажной или малоэтажной застройки, земли рекреации) могут быть использованы ячейки размером 10 x 10 м., а для отображения почв – 50 x 50 м., для управления землепользованием области 5 x 5 км.

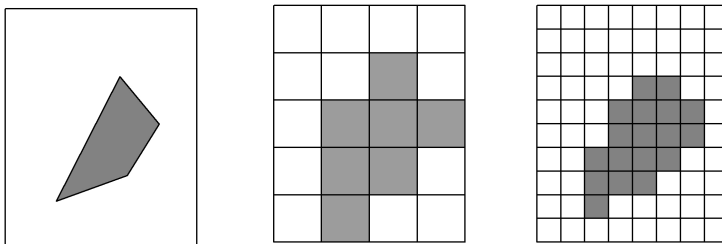


Рис. 2.4.7 - Отображение географического объекта разным пространственным разрешением

При необходимости оперировать различным пространственным разрешением применяются системы вложенных друг в друга территориальных ячеек.

В общем случае разрешение выражают в следующих единицах:

- ppi (pixel per inch) - пиксел на дюйм,
- dpi (dot per inch) - точек на дюйм,
- lpi (line per inch) - линий на дюйм.

2.4.2.2 Геометрия растров

Форма, размер, количество ячеек растра

Растр имеет прямоугольную форму. Каждая ячейка имеет прямоугольную форму и размеры: ширину d_x и высоту d_y .

Растр образован ячейками из p рядов и m колонок. Общее количество ячеек растра будет $p \cdot m$. Растр имеет $p \cdot m$ центральных точек, $(p+1) \cdot (m+1)$ точек углов. Это различие между количеством центральных точек и количеством углов часто игнорируется, и о растре просто думают как множество ячеек.

Растровые ряды и колонки ориентированы по осям координат растра.

Ориентирование растра

Ориентирование определяется углом A поворота растра относительно выбранной системы координат в реальном земном пространстве – геодезической системы координат.

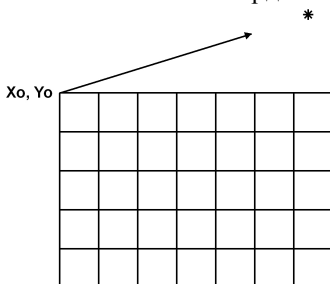


Рис. 2.4.8 – Ориентирование растра.

2.6.2.3 Координаты ячеек

Система координат растра

В системе координат растра для каждого растра вводится:

- или левая система координат, если за начало отсчета принимают верхний левый угол
- или правая система координат, если за начало отсчета принимают нижний левый угол матрицы.

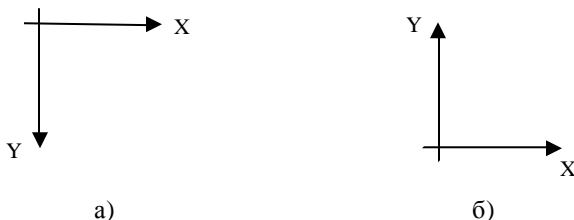


Рис. 2.4.9 - Система координат а) левая, б) правая.

В растровых моделях должна быть определенность, к какой точке пространственной ячейки относятся координаты - к одному из углов ячейки или к центральной точке ячейки.

Если за начало растра принимают его верхний левый угол с координатами x_0, y_0 , в этом случае угловые точки растра имеют координаты растра:

- нижняя левая точка:
 $x_0, \quad y_0 + n * d_y;$
- верхняя правая точка:
 $x_0 + m * d_x, \quad y_0;$
- нижняя правая точка:
 $x_0 + m * d_x, \quad y_0 + n * d_y .$

Верхний левый угол ячейки в колонке i и ряду j , имеет координаты растра:

$$x_i = x_0 + (i - 1) * d_x \quad (2.4.1)$$

$$y_j = y_0 + (j - 1) * d_y \quad (2.4.2)$$

Центральная точка ячейки в колонке i и ряду j имеет координаты растра:

$$x_i = x_0 + (i - 0,5) * d_x \quad (2.4.3)$$

$$y_j = y_0 + (j - 0,5) * d_y \quad (2.4.4)$$

Геодезическая система координат

Для определения координат ячеек в геодезической системе координат должны быть получены геодезические координаты верхнего левого угла растра X_0, Y_0 .

При угле ориентирования A растра геодезические координаты верхнего левого угла ячейки в ряду i и колонке j вычисляются по формулам преобразования координат:

$$X_i = X_0 + x_i * \cos A - y_j * \sin A \quad (2.4.5)$$

$$Y_j = Y_0 + x_i * \sin A + y_j * \cos A \quad (2.4.6)$$

2.4.2.4 Экстент региона

Для определения объема данных может быть использован экстенд моделируемого региона. При этом число колонок и рядов определяется в 2 шага:

- 1) определение размеров региона по направлениям координатных осей по формулам:

$$\Delta X_{\min, \max} = X_{\max} - X_{\min} \quad (2.4.7)$$

$$\Delta Y_{\min, \max} = Y_{\max} - Y_{\min} \quad (2.4.8)$$

- 2) вычисление требуемого числа n рядов и m колонок по формулам:

$$m = \Delta X_{\min, \max} / dx \quad (2.4.9)$$

$$n = \Delta Y_{\min, \max} / dy \quad (2.4.10)$$

2.4.2.5 Топология ячеек растров

Все ячейки, не ограничивающие растр, имеют 4 соседних ячейки. Эти четыре соседних ячейки все совместно используют край ячейки. Если в соседние ячейки включены диагональные ячейки, общее количество ячеек будет 8. Таким образом, количество ячеек с полным набором 4 или 8 соседей равно:

$$k = (n-2)*(m-2) \quad (2.4.11)$$

Кроме этого, есть ячейки на границе, которые имеют только трех соседей. Количество таких ячеек равно:

$$k = 2*(n-2) + 2*(m-2) \quad (2.4.12)$$

Наконец, в углах растра есть 4 ячейки только с 2 соседними ячейками.

На основании изложенного общее количество ячеек с разным количеством соседей равно:

$$k = (n-2)*(m-2) + 2*(n-2) + 2*(m-2) + 4 = n*m \quad (2.4.13)$$

2.4.2.6 Значение ячеек растра

Каждая ячейка растра описывается тремя параметрами:

- 1) номер S колонки (Column),
- 2) номер R ряда (Row),
- 3) значение V ячейки (Value).

Эти параметры образуют позиционную и содержательную часть.

Позиционная часть представляется номером строки и номером столбца.

Содержательная часть представляется смысловым кодом - значением, с которым может быть связан неограниченный набор атрибутов. *Значение* – это элемент информации, хранящийся в ячейке растра. Как правило, каждая ячейка растра или сетки (грида) должна иметь лишь одно значение. Ячейкой фиксируется элементарный объект поверхности однородного (гомогенного) свойства. В случае,

когда граница двух типов покрытий может проходить через часть элемента растра, значение ячейки может быть определено как:

- среднее для всех значений в ячейке,
- преобладающее значение из всех значений в ячейке,
- значение в центре ячейки,
- значение в угле ячейки.

2.4.2.7 Кодирование значений

Информацию ячейки могут представлять целые, действительные, буквенные значения. Во многих случаях для обработки информации эти значения целесообразно представлять кодами. Кодирование – это присвоение каждой ячейки кода ее атрибутов. В ГИС применяют различные приемы кодирования информации.

0	0	1	1	1	1
0	0	0	1	1	1
0	2	2	2	1	1
2	2	2	2	1	3
3	2	3	3	3	3
3	3	3	3	3	3

Рис. 2.4.10 - Кодирование ячеек

2.4.2.8 Цветовые модели

В качестве примера рассмотрим кодирование цветового пространства.

Значения пространственных ячеек растров, определенных по аэрокосмическим снимкам, – это цветовые характеристики растра. Они несут богатую цветовую информативность. Цвета пространственных ячеек растров могут быть представлены в цифровом виде посредством определенной системы представления цветов.

Цветовое пространство непрерывно. Любой цвет может быть получен в результате синтеза трех исходных базовых цветов.

Первичными базовыми цветами являются :

- красный (R - Red);
- зеленый (G - Green);
- синий (B - Blue);

вторичными базовыми цветами являются :

- голубой (C - Cyan);
- пурпурный (M - Magenta);
- желтый (Y - Yellow);

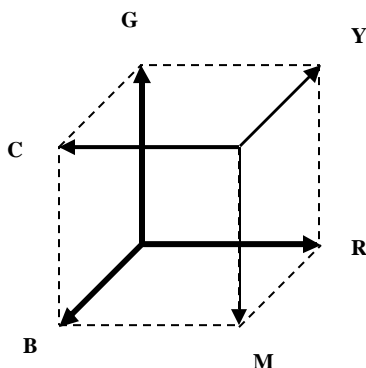


Рис. 2.4.11 – Модель пространства цветов

Для цифрового представления цветов используют интенсивность каждого цвета, которая может принимать 256 дискретных значений ($256 = 2^8$) от 0 до 255. На основании интенсивности базовых цветов, созданы системы представления цветов:

- модель *RGB*, по которой интенсивность каждого цвета определяется по формуле:

$$c * C = r * R + g * G + b * B, \quad (2.4.14)$$

где r, g, b – интенсивность цветов Red (R), Green (G), Blue (B) соответственно;

- модель *CMY*, по которой интенсивность каждого цвета определяется по формуле:

$$a * A = c * C + m * M + y * Y, \quad (2.4.15)$$

где c, m, y – интенсивность цветов Cyan (C), Magenta (M), Yellow (Y) соответственно.

- модели *CMYK* (+ black), *HSB*, *Grayscale* и др.

Количество цветов модели равно

$$256 * 256 * 256 = 16,7 \text{ млн.}$$

Чем меньше интенсивность, тем темнее цвет. В модели *RGB* интенсивность всех 3 каналов равная 0 образует черный цвет

$$c * C = o * R + o * G + o * B, \quad (2.4.16)$$

а интенсивность всех 3 каналов равная 255 образует белый цвет

$$c * C = 256 * R + 256 * G + 256 * B. \quad (2.4.17)$$

В модели RGB, например, рекреациям можно присвоить следующую интенсивность 3 каналов

$$c * C = 211 * R + 252 * G + 190 * B. \quad (2.4.18)$$

Таким образом, каждый цвет можно представить в цифровом виде в выбранной цветовой системе.

2.6.2.9 Зонирование

В растровой модели зона – это целостный объект, который определяется множеством смежных ячеек, имеющих одинаковые значения. Во многих случаях зонирование выполняют посредством классификации значений (табл.2.4.1)

Таблица 2.4.1 – Кодирование зон стоимости

Зона	ИД	Площадь	Периметр	Стоимость
А	1	17	35	325
В	2	24	78	125
А	3	12	23	26
С	4	8	15	30
В	5	20	54	890
Д	6	18	61	207

Стоимость	Код зоны стоимости
0-100	0
101-500	1
501-1000	2
1001-1500	3

Пример содержания отдельного слоя: результат одной полосы ДЗЗ со спутника дает уровень радиации, часто зарегистрированный как число между 0 и 255 (8 битов). Классифицируемая сцена обозначает различные типы использования земли, например: 1 = застроенные земли, 2 = культивируемая земля, 3 = вода.

Следует обратить внимание на то, что любой образец битов может использоваться, чтобы обозначить класс. Эти образцы битов могли бы соответствовать числам, или буквам, или знакам в схеме кодирования знаков компьютера, например * = застроенные земли, & = культивируемая земля, \$ = вода.

При наличии разных наборов значений одних и тех же ячеек матрицы формируют разные растры. База данных может содержать много таких слоев. Для совместного их использования, требуется:

- пространственное совпадение каждого слоя со всеми другими слоями,
- идентичные числа рядов и колонок,
- идентичные местоположения в плане.

2.4.3 Хранение растровых данных

Растры применяются для хранения и обработки данных дистанционного зондирования, для создания и анализа цифровых моделей поверхности, при визуализации геоданных и т.д. Для хранения растровых данных разработано множество вариантов кодирования растровых структур. Некоторые из них более экономно расходуют память, другие позволяют получать более быстрые алгоритмы.

Способ полного представления (Exhaustive representation) – это прямой ввод одной ячейки за другой. Растровое изображение обычно разлагается по строке сверху – слева. При этом порядке сканирования в конце каждой строки происходит скачок на начало следующей строки. Таким образом, двумерное ячеестое изображение хранится в памяти компьютера в виде одномерной последовательности значений. Это самый простой способ ввода данных растровых моделей.

1	1	4	4	4	4
2	2	2	1	1	1
2	2	3	1	1	2
3	3	3	3	2	2
3	1	1	1	4	2
1	1	1	4	4	4
1	1	4	4	4	4

Ячейка	Значение
1	1
2	1
3	4
4	4
5	4
6	4
7	2
8	2
9	2
10	1
11	1
12	1
13	2
14	2
15	3
16	1
17	1
18	2
19	3
...	...
21	4

Рис. 2.4.14 - Способ полного представления

Способ последовательного кодирования (Run-length encoding) основан на фиксировании повторяющегося значения и позиции – номера колонки с последним этим значением.

1	1	4	4	4	4
2	2	2	1	1	1
2	2	3	1	1	2
3	3	3	3	2	2
3	1	1	1	4	2
1	1	1	4	4	4
1	1	4	4	4	4

Значение	Позиция	Значение	Позиция	Значение	Позиция	Значение	Позиция
1	2	4	6				
2	3	1	6				
2	2	3	3	1	5	2	6
3	4	2	6				
3	1	1	4	4	5	2	6
1	3	4	6				
1	2	4	6				

Рис. 2.4.15 – Способ последовательного кодирования

2.4.4. Сжатие растровых данных

Самый простой способ ввода данных растровых моделей - прямой ввод одной ячейки за другой. Недостатками этого подхода являются требования большого объема памяти в компьютере и значительного времени для организации процедур ввода-вывода. Например, снимок искусственного спутника Земли Lansat имеет 74 000 000 элементов растра. Это требует огромных ресурсов для хранения данных.

При растровом вводе информации в ГИС возникает проблема ее сжатия, так как наряду с полезной информацией может попадать и избыточная (в том числе и бесполезная) информация.

Сжатие данных (Data compression) - процесс, обеспечивающий уменьшение объема данных путем сокращения их избыточности. Сжатие данных связано с компактным расположением порций данных стандартного размера. Различают сжатия с потерей и без потери информации.

2.4.4.1 Метод группового кодирования.

В некоторых форматах графических файлов используется сжатие изображения, основанное на замене длительных последовательностей повторяющихся значений парой - "значение, количество повторов".

Для сжатия информации, полученной со снимка или карты, применяется кодирование участков развертки или метод группового кодирования, учитывающий то, что значения повторяются в нескольких ячейках. Суть метода группового кодирования состоит 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 5x6 (Рис. 2.6.16):

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

Рис. 2.4.16 – Массив значений

При использовании метода группового кодирования он будет вводиться следующим образом:

30312031303120511051

Вместо 30 необходимо только 20 элементов данных. В рассмотренном примере коэффициент сжатия составляет 1:1,3. Экономия составляет 30 %. На практике при работе с большими массивами бинарных данных требуется гораздо большее сжатие.

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 раstra 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.

2.4.4.2 Методы, основанные на порядке сканирования

Географические данные обычно автокоррелированы. Это означает, что соседние ячейки растровой модели имеют большую вероятность быть одинаковыми, чем разобщенные, а географические объекты образуют на растровом изображении пятна.

Способ двунаправленного сканирования строк (способ Boustrophedon – по названию быка, вспахивающего поле) основан на изменении порядка кодирования: нечетные строки кодируются слева направо, а четные – в обратном направлении. Теперь при переходе к новой строке первая ячейка является смежной последней ячейке старой строки.

A	A	A	A
A	B	B	B
A	A	B	B
A	A	A	B

Рис. 2.4.17 – Матрица сканирования

Порядок сканирования раstra способом полного представления:

AAAA ABVV AABV AAAV (16 байт);

5A 3B 2A 2B 3A 1B (12 байт).

Порядок сканирования раstra способом двунаправленного кодирования:

AAAA BVVA AABV BVAAA (16 байт);

4A 3B 3A 4B 3A (8 байт).

В линейном разложении раstra автокорреляция значений ячеек учитывается только по одному направлению (по строке). При этом повышается эффективность кодирования.

С целью охвата линией обхода двумерных пятен растровой модели разработан ряд приемов сканирования ячеек. Например, порядок сканирования Мортон (Morton, впервые использовавшего этот способ в Canada GIS) основан на иерархическом разбиении карты. При сканировании раstra по Мортону линия сканирования представляет собой фрактал (геометрическая фигура, части которой подобны ей самой). Недостатками этого порядка сканирования являются присутствие скачков между ячейками и возможность кодировать

только растры размера, кратного двум. В порядке сканирования Пеано отсутствуют скачки между ячейками. Ломаная линия Дж. Пеано (1890) имеет базовый П-образный шаблон, который поворачивается от уровня к уровню так, чтобы обеспечить непрерывность линии сканирования.

2.4.4.3 Квадратомическое дерево

Квадратомическое представление (Quadtree) - один из способов представления пространственных объектов в виде иерархической древовидной структуры, основанный на декомпозиции пространства на квадратные участки, или квадратные блоки, квадранты (quarters, quads), каждый из которых делится рекурсивно на 4 вложенных до достижения некоторого уровня - числа Мортон (Morton order), обеспечивающего требуемую детальность описания объектов, эквивалентную разрешению растра [16].

Квадратомическое дерево обычно используется как средство снижения времени доступа, повышения эффективности обработки и компактности хранимых данных по сравнению с растровыми представлениями.

Квадратомическая структура описывается иерархической системой вложенных квадратов.

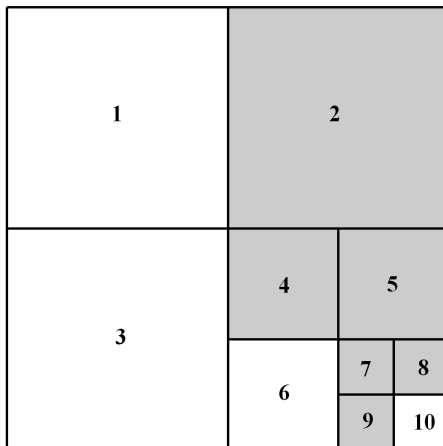


Рис. 2.4.18 - Представление области иерархической древовидной структурой вложенных квадратов

На каждом уровне выделяются квадраты:

- - заполненные,
- ◼ - частично заполненные,
- - незаполненные.

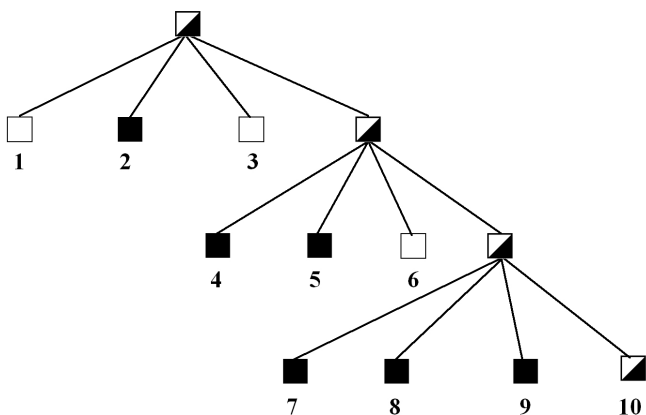


Рис. 2.4.19 - Иерархическое представление области в виде дерева вложенных квадратов

Записи атрибутов квадротомического дерева представлены в табл. 2.4.2.

Табл. 2.4.2 - Атрибуты квадротомического дерева

Уровень	Уровень	Уровень	Значение
1	2	3	
1			Незаполненный
2			Заполненный
3			Незаполненный
	4		Заполненный
	5		Заполненный
	6		Незаполненный
		7	Заполненный
		8	Заполненный
		9	Заполненный
		10	Частично заполненный

В этом примере коэффициент сжатия составляет 1:6,4 (10 записей вместо 64).

Основное преимущество иерархической организации данных на quadroдеревьях заключается в том, что она позволяет получать быстрые способы доступа к пространственным данным.

2.4.5 Форматы растровых данных

Формат (format) - это способ расположения или представления данных в памяти, базе данных, документе или на внешнем носителе. В ГИС, машинной графике и обработке изображений формат - это общее наименование способа машинной реализации представления (модели) пространственных данных или формат данных конкретной системы, программного средства, средства стандартизации, обмена данными.

Для хранения растровых данных используются многие форматы; некоторые из них представлены в табл.2.4.3.

Таблица 2.4.3 – Форматы растровых данных

Название формата	Характеристика формата
BMP (Microsoft Windows/IBM® OS/2® Bitmap)	Битовый двоичный формат растровых графических файлов, предусматривающий 4, 8 и 24 бита на точку.
TIFF (Tagged Image File Format)	Формат для создания и обмена изображениями высокого качества; поддерживает цветные (до 24 бит), черно-белые изображения и градации серого; использует разные типы сжатия данных с учетом фотометрических свойств изображения.
TIFF-LZW	TIFF-LZW ERDAS Imagine - формат сжатия растровых данных, основанный на алгоритме Lempel Ziff & Welsh, поддерживается при наличие специальной библиотеки. Коэффициент сжатия изображения, представленного в формате TIFF, составляет 1:7,7
GeoTIFF	Расширение формата TIFF для передачи изображений, имеющих пространственную привязку, включает информацию о системе координат и проекции, параметрах геометрической коррекции.

Продолжение таблицы 2.4.3.

JPEG (Joint Photographic Experts Group)	Формат для сжатия изображений с коэффициентом более 25:1 с потерей качества; позволяет передавать до 16 млн. цветов (до 32 бит); распространен для изображений в Интернете.
JFIF (JPEG File Interchange Format)	Формат файлов для хранения и передачи изображений, сжатых по алгоритму JPEG.
PCX	Распространенный формат графических файлов для цветных изображений, используемый в большинстве графических редакторов. Изображение сжимается до 1:1.5.
GIF (Graphics Interchange File)	Способ сжатия графического изображения. Формат графических файлов, наиболее часто используемый в Интернете; обеспечивает сжатие данных без потерь; алгоритм сжатия LZW
ERDAS IMAGE	Файлы разработаны, используя программное обеспечение обработки изображений IMAGE. Файлы ERDAS IMAGE могут хранить непрерывные и дискретные одноканальные или многоканальные данные.
IMG (Image)	Формат сканированных с высоким разрешением изображений. Графический файл в пакете GEM. Используется во многих растровых ГИС-пакетах, например, ERDAS.
ER Mapper	ER Mapper файлы разработаны, используя программное обеспечение обработки изображений ER Mapper.
ESRI Grid	Формат, разработанный фирмой ESRI, поддерживающий 32-битовые целые и действительные растровые сетки. Целые гриды предназначены для представления дискретных данных, действительные гриды используются для представления непрерывных явлений.
MrSID™ (Multiresolution Seamless Image Database)	Формат с переменной разрешающей способностью на основе метода волновой модуляции, который обеспечивает высокую степень сжатия.

Из таблицы следует, что проблема ограниченного дискового пространства может быть решена путем использования компрессированных растров, что приведет к сокращению объемов файлов. Одним из самых эффективных алгоритмов сжатия является алгоритм Lempel-Zip-Welch (LZW), применяемый в формате GIF, LZW. Компрессия также поддерживается в форматах TIFF и IMG.

При этом важно знать о возможных потерях информации. Примером формата сжатия с потерями информации может быть JPEG; примерами форматов сжатия без потери информации могут быть GIF, TIFF.

Разработаны форматы, например, MrSID, дающие большой коэффициент сжатия.

2.4.7 Файл геопривязки растровых данных

Файлы растровых данных в выбранном формате представляют изображение, в котором положение каждой ячейки фиксируется координатами изображения - номером ряда и номером колонки. Эти данные не содержат информацию о координатах ячеек в принятой системе координат для реального земного пространства, и поэтому их недостаточно для представления географических объектов.

Для представления географических объектов необходимо установить преобразование координат изображения в координаты реального мира. Специализированные форматы, такие как ERDAS IMAGE, BSQ, BIL, BIP, GeoTIFF, Grid хранят геопространственную информацию в заголовке файла изображения. Растровые данные, представленные в общих форматах, необходимо совместно использовать с отдельными файлами геопривязки - World файлами, которые содержат информацию о преобразовании координат изображения в координаты реального мира. Файл геопривязки имеет тоже название, что и файл изображения, но его расширение дополнено в конце буквой **w** (от World).

Таблица 2.4.4 – Примеры файлов геопривязки

Файл изображения	Файл геопривязки
zones.tif	zones.tifw или zones.tifw
cost.jpg	cost.jpgw или cost.jpgw
image.bil	image.bilw или image.bilw

Файл геопривязки представляется в формате ASCII; он содержит информацию для преобразования координат изображения в координаты реального земного пространства (табл.2.4.5).

Таблица 2.4.5 – Пример содержания файла геопривязки

Аргументы	Пример значений аргументов
Размер пиксела по оси X	0.423566666650000
Коэффициенты поворота	0.000000000000000 0.000000000000000
Размер пиксела по оси -Y	-0.423566666650000
Координаты центра верхнего левого пиксела	79000.000000000000000 71000.000000000000000

При наличии файла геопривязки программа выполняет трансформацию координат по формулам аффинных преобразований:

$$x' = Ax + By + C \quad (2.4.19)$$

$$y' = Dx + Ey + F \quad (2.4.20)$$

где:

- x, y – вычисляемые координаты i -пиксела в принятой системе координат для реального земного пространства,
- x, y – номер колонки, номер ряда пиксела на карте соответственно,
- A, E – размер пиксела соответственно по оси X, по оси -Y,
- B, D – коэффициенты поворота,
- C, F – координаты x_0, y_0 центра верхнего левого пиксела в принятой системе координат для реального земного пространства.

2.4.8 Резюме представления географических объектов растровыми моделями

Обобщенная характеристика представления географических объектов растровыми моделями приведена в табл.

Таблица 2.4.5 – Резюме представления географических объектов растровыми моделями [29].

Фокус модели	Триангуляционные данные ориентированы на моделирование непрерывных явлений и изображений земной поверхности.
Источники данных	Материалы аэрофотосъемки, космической съемки сбор GPS определений, конвертирование из триангуляционных данных, растеризация векторных данных, Сканирование фотографий и чертежей.
Хранение пространственных данных	От координат нижнего левого угла раstra и высоты, ширины ячейки, каждая ячейка локализована позицией ее ряда и колонки.
Представление пространственных объектов	Точечные пространственные объекты представляются одной ячейкой. Линейные пространственные объекты представляются цепочкой смежных ячеек с одинаковым значением. Полигональные пространственные объекты представляются регионами ячеек с одинаковым значением.
Топологические отношения	Смежные ячейки могут быстро локализованы приращением и вычитанием значений рядов и колонок.
Геопространственный анализ	Пространственное совпадение, анализ близости, анализ поверхности, дисперсия, путь наименьших затрат.
Картографическая продукция	Растровые данные являются лучшими для презентации изображений и непрерывных явлений с постепенным варьированием атрибутов. Они не совсем подходят для вычерчивания точечных и линейных пространственных объектов.

2.4.8 Контрольные вопросы и задания для самостоятельной работы

- 1) В чем состоит сущность растровых моделей географических объектов?
- 2) Как представляются дискретные географические объекты в растровых моделях?
- 3) Дайте определение поверхности.
- 4) Как представляются непрерывные явления растровыми и сеточными моделями?
- 5) В чем состоят преимущества и недостатки растровых моделей?
- 6) Чем определяется разрешение растра?
- 7) Как определяются координаты ячеек растра в системе координат растра и геодезической системе координат?
- 8) Чем определяется экстенс региона, топология ячеек растра?
- 9) Опишите известные вам цветовые модели.
- 10) Что есть зонирование? Как можно использовать таблицу атрибутов растра для операций зонирования?
- 11) Опишите способ полного представления и способ последовательного кодирования для хранения растровых данных.
- 12) Опишите сущность метода группового кодирования и методов, основанных на порядке сканирования для сжатия растровых данных.
- 13) Опишите сущность метода квадротомического дерева для сжатия растровых данных.
- 14) Приведите краткую характеристику форматов хранения растровых данных.
- 15) Что представляет собой файл привязки растровых данных?
- 16) Приведите обобщенную характеристику представления географических объектов растровыми моделями.

Раздел 2.5

ТРИАНГУЛЯЦИОННЫЕ МОДЕЛИ ГЕОГРАФИЧЕСКИХ ОБЪЕКТОВ

2.5.1 Определение модели TIN

Модель TIN (Triangulation Irregular Network - триангуляционная нерегулярная сеть) географических объектов – модель поверхности в виде сети смежных не пересекающихся треугольных граней, определенная по узлам и ребрам, которые покрывают поверхность.

Слова в названии модели означают следующее:

- Triangulation (Триангуляционная) – указывает способ построения оптимизированного набора треугольников по набору точек;
- Irregular (Нерегулярная) – точки могут быть взяты с переменной плотностью для моделирования участков поверхности, где рельеф резко меняется;
- Network (Сеть) – отражает топологическую структуру.

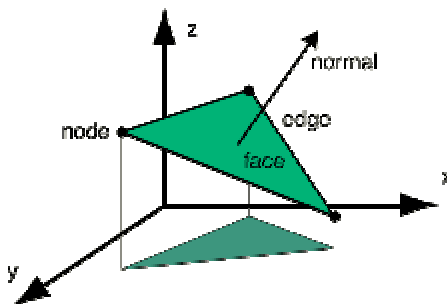


Рис. 2.5.1 – Геометрия модели TIN [29]

Геометрия модели TIN образуется гранями, узлами и ребрами в трехмерном пространстве.

- Грань (Face) – поверхность треугольника в трехмерном пространстве.
- Узел (Node) – вершина треугольника с координатами X,Y,Z.
- Ребро (Edge) – сторона треугольника в трехмерном пространстве.

Каждая грань TIN является частью поверхности в 3D-пространстве.

Модель TIN является альтернативой растровым DEM и используются во многих геоинформационных системах, системах автоматизированного картографирования. Модели TIN разработаны в 1970-х годах как простой способ построения поверхностей по нерегулярно расположенным точкам.

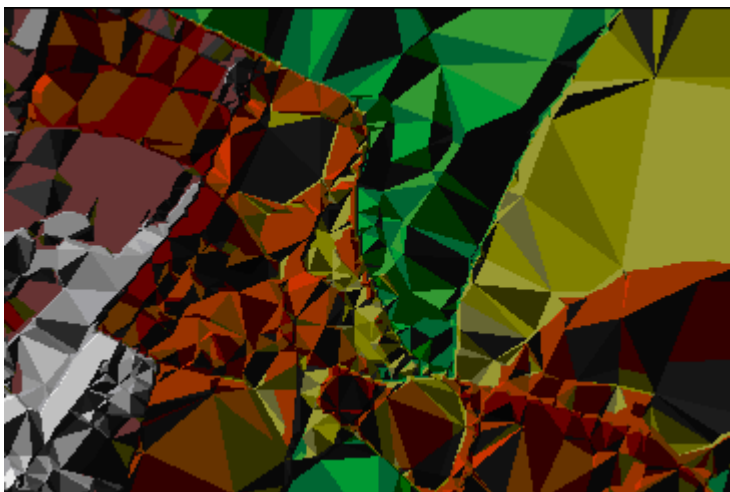


Рис. 2.5.2 – Пример отображения фрагмента поверхности стоимости земли моделью TIN

2.5.2 Свойства модели TIN

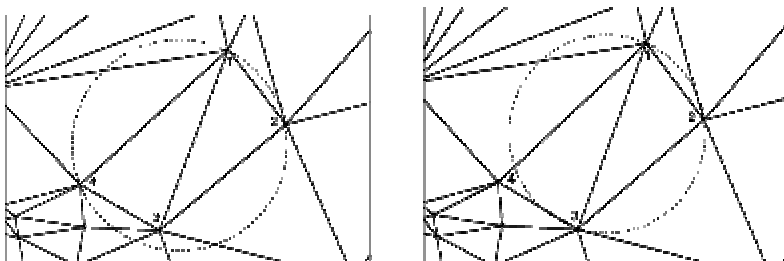
Модель TIN обладает следующими свойствами:

- Модель TIN позволяет получить точное представление о локальной части поверхности, используя переменную плотность узлов со значением Z и линии перегиба поверхности;
- Модель TIN является основой 3D-визуализации поверхности;
- Модель TIN позволяет выполнить сложный анализ поверхности (вычисление высот, уклонов, экспозиций склонов, получение изолиний поверхности, расчеты объемов, вертикальные профили по трассе линии, анализ видимости).

2.5.3 Триангуляция Делоне

Исходными данными для построения TIN является набор точек с координатами X, Y, Z . Задача заключается в том, чтобы по этому набору точек создать сеть смежных непересекающихся треугольников. Задача построения триангуляции по набору точек является одной из базовых в вычислительной геометрии. К ней сводятся многие другие задачи, она широко используется в машинной графике и геоинформационных системах для моделирования поверхностей и решения пространственных задач.

Задача построения триангуляции по исходному набору точек является неоднозначной, поэтому возникает вопрос, какая из различных триангуляций лучше? Можно, например, оптимальным решением считать такое, при котором сумма длин всех рёбер будет минимальной среди всех возможных триангуляций, построенных на тех же исходных точках. Решение задачи при таком условии имеет высокую трудоёмкость [34].



а) Формирование треугольника в триангуляции Делоне: а) круг, построенный по точкам 1,2,4, включает точку 3; б) круг, построенный по точкам 1,2,3, не включает точку 4.

Наибольшее распространение в ГИС получила триангуляция Делоне (Delaunay), которая названа по имени ее автора советского математика Бориса Николаевича Делоне (1890-1980). По определению Делоне три точки формируют треугольник в триангуляции тогда и только тогда, когда в окружности, описанной вокруг этого треугольника нет других точек разбиения. Каждый ограничивающий треугольник круг не содержит точек из набора внутри его (рис.2.5.3,б).

Один из алгоритмов построения триангуляции Делоне основан на генерировании полигонов Тиссена (Thiessen) или Вороного. Для этого поверхность разбивается на области, в которых каждая точка

расположена ближе всего к некоторому узлу сети – генерирующей точке. Полученные границы называют полигонами Тиссена или полигонами Вороного. Две точки соединяются линией в триангуляции Делоне, если их полигоны Тиссена имеют общую границу. Этот метод позволяет получить требуемые треугольники. Полигоны Тиссена используются также при анализе близости.

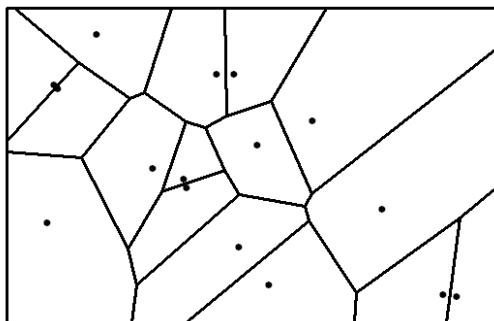


Рис. 2.5.4 - Полигоны Тиссена

2.5.4 Топология в TIN

Модель TIN – это топологическая структура данных: ребра соединяются в узлах; каждый треугольник смежный с соседними треугольниками.

Топологические отношения создаются путем создания в базе данных для каждого узла указания на смежные узлы. Пространство вокруг территории представляется фиктивным узлом.

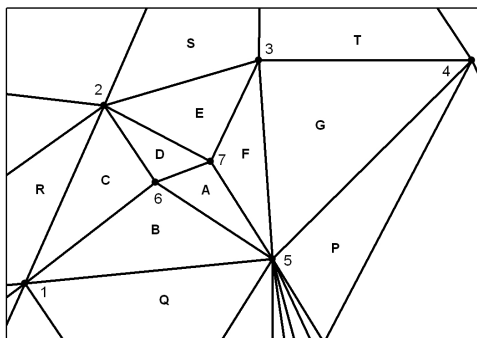


Рис. 2.5.5 - Модель TIN

Модель TIN (рис. 2.5.5) в цифровом виде описывается связанными файлами - файлом вершин (табл.2.5.1), файлом указателей и файлом треугольников (табл.2.5.2). Для каждой точки разбиения сохраняется ее уникальный номер, координаты и список точек, с которыми она соединена прямыми (по часовой стрелке).

Табл.2.5.1 – Список вершин

Вершина	X	Y	Z
1	X_1	Y_2	Z_3
2	X_2	Y_2	Z_2
3	X_3	Y_3	Z_3
4	X_4	Y_4	Z_4
5	X_5	Y_5	Z_5
6	X_6	Y_6	Z_6
7	X_7	Y_7	Z_7

Табл.2.5.2 – Список треугольников

Треугольник	Вершины			Смежные треугольники		
	5	6	7	B	D	F
A	5	6	7	B	D	F
B	1	5	6	A	C	Q
C	1	2	6	B	D	R
D	2	6	7	A	C	E
E	2	3	7	D	F	S
F	3	5	7	A	E	G
G	3	4	5	F	T	P

2.5.5 Этапы создания модели TIN

Модель TIN создается в следующей последовательности.

Этап 1. Задание множества точек i по координатам X_i, Y_i, Z_i .

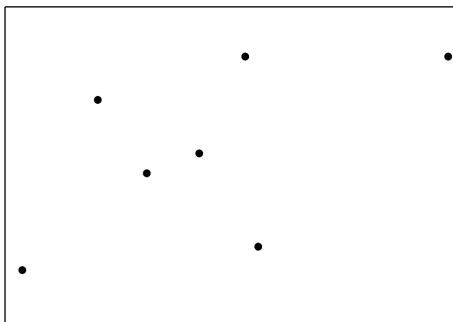


Рис. 2.5.6 - Множество точек

Этап 2. Построение триангуляция Делоне

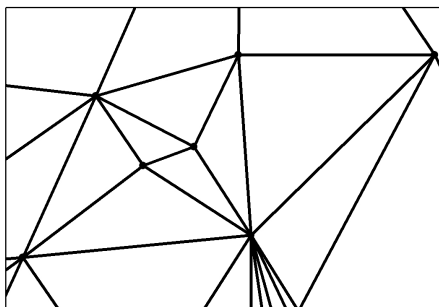


Рис. 2.5.7 – Триангуляция Делоне

Этап 3. Ввод линий перегиба поверхности (Breakline) и модификация TIN с учетом линий перегиба. Линии перегиба рельефа определяют резкие изменения поверхности, такие как линия верха, низа откоса, гребни, тальвеги и др.

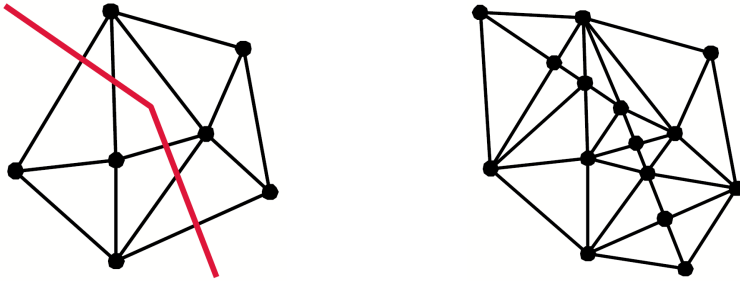


Рис. 2.5.8 – Линии перегиба поверхности [29]

Этап 4. Ввод областей исключения (Exclusion areas) с постоянным значением Z и модификация TIN с учетом полигональных объектов, например водной поверхности.

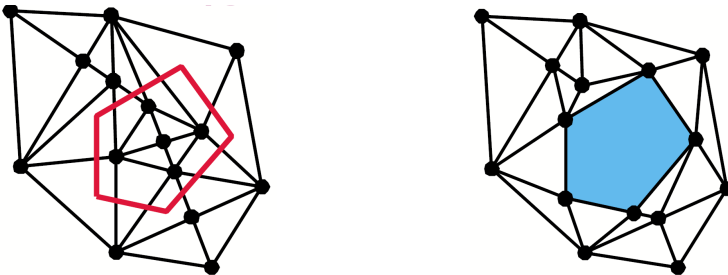


Рис. 2.5.9 - Области исключения [29]

Этап 5. Решение задач по модели TIN поверхности.

2.5.6 Уравнение нормали к грани треугольника

В TIN-модели пространственное положение треугольника характеризует нормаль к его грани.

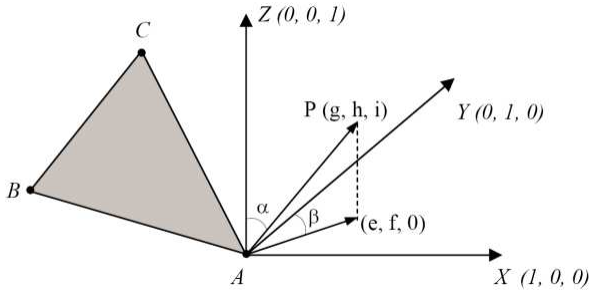


Рис.2.5.10 – Нормаль P к грани треугольника

Треугольник задается тремя точками (x_a, y_a) , (x_b, y_b) , (x_c, y_c) . Нормаль P к плоскости треугольника определяется нормальным вектором $P = \{p_a, p_b, p_c\}$, который получается в результате векторного произведения векторов AB и AC двух его сторон. Для нормального вектора P каждого треугольника вычисляются:

1) прямоугольные декартовы координаты:

$$\begin{aligned} p_a &= [(y_b - y_a) * (z_c - z_a) - (z_b - z_a) * (y_c - y_a)] \\ p_b &= [(z_b - z_a) * (x_c - x_a) - (x_b - x_a) * (z_c - z_a)] \\ p_c &= [(x_b - x_a) * (y_c - y_a) - (y_b - y_a) * (x_c - x_a)] \end{aligned} \quad (2.5.1)$$

2) направляющие косинусы:

$$g = \frac{p_a}{\sqrt{p_a^2 + p_b^2 + p_c^2}} \quad (2.5.2)$$

$$h = \frac{p_b}{\sqrt{p_a^2 + p_b^2 + p_c^2}} \quad (2.5.3)$$

$$i = \frac{p_c}{\sqrt{p_a^2 + p_b^2 + p_c^2}} \quad (2.5.4)$$

3) проекции к горизонтальной плоскости:

$$e = \frac{p_a}{\sqrt{p_a^2 + p_b^2}} \quad (2.5.5)$$

$$f = \frac{P_b}{\sqrt{P_a^2 + P_b^2}} \quad (2.5.6)$$

Эти зависимости являются базовыми для решения задач отображения и анализа поверхности.

2.5.7 Средства модели TIN для отображения поверхности

Модель TIN имеет эффективные средства отображения поверхности посредством экспозиции склонов, крутизны граней, затенения граней, диапазонов высот для поверхности.

2.5.7.1 Экспозиция склона

Экспозиция (Aspect) отдельной грани определяется углом β как направлением проекции нормали на горизонтальную поверхность относительно направления на Север. Угол вычисляется по формуле:

$$\beta = \arccos(f) = \arccos\left(\frac{P_c}{\sqrt{P_a^2 + P_b^2}}\right) \quad (2.5.8)$$

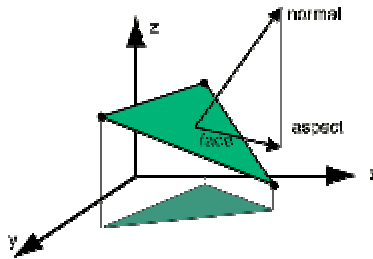


Рис. 2.5.11 – Отображение экспозиции склонов

2.5.7.2 Крутизна грани

Крутизна грани (Slope) представляется углом наклона нормали к горизонтальной линии. Он определяется как дополнение до 90^0 угла α :

$$\alpha = \arccos(i) = \arccos\left(\frac{P_c}{\sqrt{P_a^2 + P_b^2 + P_c^2}}\right) \quad (2.5.9)$$

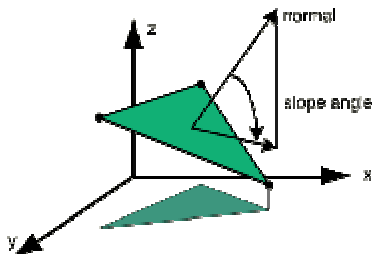


Рис. 2.5.12 – Отображение крутизны граней

2.5.7.3 Затенение граней

Затенение граней (Hillshading) определяется относительно вектора направления на светило. Положение светила задается азимутом (Azimuth) и углом возвышения над горизонтом (Altitude) светила

Задача расчета затенения склонов обычно используется для анализа освещенности Земли. В связи с этим часто возникает потребность дополнительного учета текущего положения солнца. Таким образом, каждый треугольник триангуляции может быть проклассифицирован по принципу принадлежности к тому или иному региону. После этого нужно просто вызвать алгоритм выделения регионов.

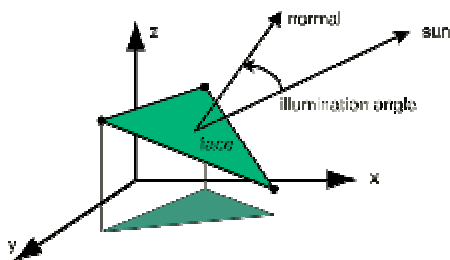


Рис. 2.5.13 – Отображение затенения граней

2.5.7.4 Диапазоны высот

Для отображения диапазонов высот (Range of elevations) высота произвольной точки внутри треугольника определяется по уравнению плоскости, заданной вершинами треугольника. Плоскость с нормальным вектором $P = \{p_a, p_b, p_c\}$, проходящая через точку $M_0(x_0, y_0, z_0)$, описывается уравнением:

$$p_a(x - x_0) + p_b(y - y_0) + p_c(z - z_0) = 0 \quad (2.5.10)$$

Отсюда по известным значениям x и y находятся высоты произвольных точек.

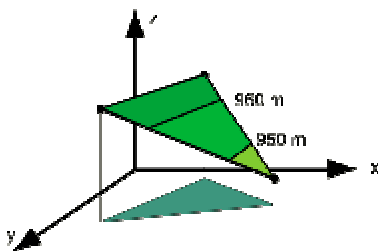


Рис. 2.5.14 – Отображение диапазонов высот

Интерполяция изолиний выполняется по каждой грани триангуляции.

2.5.8 Резюме представления географических объектов триангуляционными моделями

Обобщенная характеристика представления географических объектов триангуляционными моделями приведена в табл. 2.5.3.

Таблица 2.5.3 – Резюме представления географических объектов триангуляционными моделями [29].

Фокус модели	Триангуляционные данные ориентированы на эффективное представление поверхности, которое может отображать концентрации высот или других свойств.
Источники данных	Компиляция материалов аэрофотосъемки, сбор GPS определений, конвертирование из векторных изолиний, импорт точек со значениями z .
Хранение пространственных данных	Каждая вершина грани триангуляции имеет x, y координаты.
Представление пространственных объектов	Значения z точек определяют форму поверхности. Линии перелома определяют структурные изменения поверхности, такие как гребни или тальвеги.
Топологические отношения	Каждый треугольник связан со смежными треугольниками.
Геопространственный анализ	Вычисления высот, крутизны и направления скатов, построение изолиний (горизонталей) поверхности, вычисление объемов, построение вертикальных профилей, анализ видимости.
Картографическая продукция	Триангуляционные данные являются лучшими для богатого представления поверхности. Эти данные могут быть визуализированы, используя цвет, чтобы представить высоты, склоны, ориентировку или освещенность в трехмерной перспективе.

2.5.9 Контрольные вопросы и задания для самостоятельной работы

- 1) Что представляет собой модель TIN географических объектов? Какими свойствами обладает модель TIN?
- 2) В чем заключается принцип формирования треугольника в триангуляции Делоне?
- 3) Чем определяются топологические отношения в модели TIN?
- 4) Опишите этапы создания модели TIN.
- 5) Получите уравнение нормали к грани треугольника.
- 6) Опишите средство отображения поверхности посредством экспозиции склонов.
- 7) Опишите средство отображения поверхности посредством крутизны граней.
- 8) Опишите средство отображения поверхности посредством затенения граней.
- 9) Опишите средство отображения поверхности посредством диапазонов высот.
- 10) Приведите обобщенную характеристику представления географических объектов триангуляционными моделями.

Часть 3

**СБОР И ПОДГОТОВКА
ГЕОГРАФИЧЕСКИХ ДАННЫХ**

Раздел 3.1

СБОР И ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ГЕОГРАФИЧЕСКИХ ДАННЫХ

Основу информационного обеспечения ГИС составляют источники пространственных данных. Информационное обеспечение ГИС остается крайне трудоемким делом. Затраты на информационное обеспечение геоинформационных проектов достигают 90% от их общей стоимости.

Об объеме пространственных данных говорят следующие результаты специально проведенных исследований: до 70% всех данных, составляющих информационные ресурсы наций, регионов и ведомств, имеют пространственную привязку или могут быть координированы, получив статус пространственных.

В современном обществе геоинформация рассматривается как национальный ресурс. Она является также товаром, который создают и покупают. Геоинформация служит основой управления в тех организациях, где нужна актуальная информация, связанная с пространственными структурами. Она сочетает в себе все свойства информации и информационных продуктов в этом ее уникальность и ценность.

Сбор информации решает две задачи: получение первичной геоинформации и частичное преобразование ее в геоданные.

3.1.1 Источники географических данных

Цифровая среда существования ГИС предполагает цифровую форму обрабатываемых данных. Существует огромный диапазон видов данных, используемых геоинформационными системами. Это отражает различные цели самих систем. Поскольку ГИС может быть использована для разнообразных приложений, например, таких как городское планирование, маркетинговые исследования или археологический анализ, исходный материал может быть трудным для всесторонней инвентаризации и классификации. Даже в рамках одного ГИС-проекта ассортимент необходимых материалов может быть разным. Несмотря на то, что вид материала будет существенно

отличаться от проекта к проекту. ГИС-специалисты должны знать о некоторых из наиболее доступных источниках данных, об их характеристиках и ограничениях.

Наиболее распространенные источники пространственных данных следующие:

- топографические карты и планы аналоговые;
- топографические карты и планы цифровые;
- данные топографических съемок, измерений электронными тахеометрами и приемниками глобальной системы позиционирования;
- материалы дистанционного зондирования Земли;
- базовые наборы геопространственных данных;
- интернет-источники цифровых пространственных данных;
- общегеографические и тематические карты;
- землеустроительная документация;
- градостроительная документация;

Наиболее распространенные источники семантических данных следующие:

- таблицы;
- текстовые документы.

Топографические карты и планы аналоговые

Карты – это наиболее известные модели реального мира. Карты используются тысячи лет для того, чтобы представлять информацию о реальном мире. Карты оказались чрезвычайно полезными для многих приложений в различных областях. Концепция и дизайн карт разработаны так, чтобы интерпретировать реальный мир явлений в правильные, четкие и понятные для многих представления. Картографическими источниками являются топографические, общегеографические и тематические карты.

Топографическая карта представляет собой уменьшенное и обобщенное изображение земной поверхности, созданное по единой математической основе и оформлению, передающее размещение и свойства основных природных и социально-экономических объектов местности. Топографические карты и планы, которые традиционно создавались на бумажной основе, и представляют собой аналоговые модели. Топографические карты являются первичными данными. Топографические карты являются источником данных для производных карт, например, тематических.

Недостатком карт является то, что они ограничены в двумерных статических представлениях, и что они всегда отображаются в определенном масштабе. Карта всегда является графическим представлением на определенном уровне детализации. На карте масштаб определяет пространственное разрешение графического представления. Чем меньше масштаб, тем меньше деталей карта может показать. С другой стороны, точность представления данных ограничивает диапазон соответствующих масштабов. Выбор надлежащего масштаба карты является одним из первых и наиболее важных шагов. Карты больших физических размеров должны быть разрезаны на листы для удобства пользования. Недостатком карт является также то, что они ограничены в отображении семантики объектов из-за ограниченного места на листе карты.

Для перевода аналоговых данных в цифровую форму выполняют оцифровку (цифрование – digitizing) аналоговых топографических карт. Для этого современные ГИС-пакеты имеют встроенные инструменты. Различают оцифровку а) дигитайзером, б) по дисплею (интерактивную или автоматизированную).

Процесс оцифровки дигитайзером состоит из следующих этапов: крепление к столу листа аналоговой карты и дигитайзера, калибровка дигитайзера, получение координат точек и построение цифровых моделей географических объектов.

Широко применяемый процесс оцифровки по дисплею состоит из следующих этапов: сканирование листов аналоговых топографических карт, привязка сканированного изображения к геодезической системе координат, получение координат точек и построение цифровых моделей географических объектов.

Топографические карты и планы цифровые

С появлением компьютерных систем аналоговая картография заменяется цифровой картографией. Сегодня, когда мы говорим о картографии, мы косвенно предполагаем цифровую картографию. Использование компьютеров в создании карт является неотъемлемой частью современной картографии. Роль карты изменяется соответственно. Все чаще карты теряют свою роль в качестве хранилища данных. Эта роль переходит к базам пространственных данных. Для карты остается функция визуализации.

Цифровые топографические карты создаются путем:

- 1) оцифровки аналоговых топографических карт,

- 2) обработки материалов топографических съемок, геодезических измерений электронными тахеометрами и приемниками глобальной системы позиционирования.
- 3) обработки материалов дистанционного зондирования Земли (аэрофотоснимков, космических снимков),
- 4) комбинированным способом.

Цифровые карты создаются разными техническими и программными средствами и поэтому получаются в разных форматах. Трансформация данных из одного распространенного формата в другой сегодня не является проблемной.

В последние годы осуществлен переход топографо-геодезической службы Украины на массовое использование цифрового картографирования. Существующая технология крупномасштабного цифрового картографирования опирается на "Класифікатор інформації, яка відображається на топографічних планах масштабів 1:5000, 1:2000, 1:1000, 1:500" [35]. В ряде предприятий для цифрового картографирования используется программное обеспечение Digital/Delta с использованием материалов аэрофотосъемки.

Существующая технология цифрового картографирования ориентирована: а) на издание бумажной топографической карты, б) определенного масштаба, в) в условных обозначениях. В этой связи возникает расхождение в представлении пространственных данных цифрового картографирования с моделью данных в геоинформационной системе. Поэтому данные цифрового картографирования нуждаются в инструментах трансформации их в информационную среду геоинформационной системы.

Данные топографических съемок, измерений электронными тахеометрами и приемниками глобальной системы позиционирования

В базу данных ГИС теперь можно загружать трехмерные данные любой инструментальной съемки, измерений электронными тахеометрами, GPS-приемниками и выполнять в ней все традиционные вычисления, необходимые для корректировки этих данных и создания координатных точек с известными уровнями ошибки. Для получения наилучшего значения точки можно произвести корректировку разрозненных данных методом наименьших квадратов.

Измерения координатной геометрии (COGO) могут быть также объединены с данными съемки, их можно обрабатывать одним и тем же способом. В базу данных геосъемки также можно включить

координаты, полученные GPS-приемниками. Обработка данных съемки и вычисления выполняются в базе геоданных в том же координатном пространстве, что и обработка других данных.

Процесс передачи данных от полевых станций, обработки данных инструментальной съемки, последующей передачи обработанных данных в системы СОГО и САПР для отрисовки и структуризации и затем передачи в ГИС для интеграции с другими данными сегодня может быть реализован в одной базе геоданных в едином координатном пространстве.

Новая возможность, существенная с точки зрения обеспечения точности базы данных ГИС, — интеграция данных геодезической съемки с местоположением ГИС-объектов на карте. Может быть установлена привязка между координатами, полученными из данных съемки, и точками на объектах. После этого объекты можно переместить в правильное положение и в таком виде сохранить их в базе данных. Можно выбрать допуски замыкания, алгоритмы конфигурирования и пакетную обработку по уравниванию. Можно добавить целиком новые ГИС-объекты, определенные по данным геодезических измерений, а также повысить точность существующих ГИС-данных. Ошибки местоположений новых объектов могут быть представлены эллипсами погрешностей. Теперь можно провести количественное сравнение допусков относительной и абсолютной ошибки, указанных в описаниях информационных продуктов, с точностью размещения объектов.

Материалы дистанционного зондирования Земли

Дистанционное зондирование Земли (ДЗЗ) — наблюдение поверхности Земли авиационными и космическими средствами, оснащёнными различными видами съёмочной аппаратуры. Рабочий диапазон съёмочной аппаратуры составляет от долей микрометра (видимое оптическое излучение) до метров (радиоволны). Методы зондирования могут быть пассивные, то есть использующие естественное отраженное или вторичное тепловое излучение объектов на поверхности Земли, обусловленное солнечной активностью, и активные, использующие вынужденное излучение объектов, инициированное искусственным источником направленного действия. Материалы дистанционного зондирования Земли, являются одним из основных источников данных для геоинформационных систем. Это, прежде всего материалы аэрофотосъемки и космической съемки.

Аэрофотоснимки

Аэрофотосъемка выполняется в основном для топографического картографирования территории, также применяются в геологии, в лесном и сельском хозяйстве.

Современные методы обработки материалов аэрофотосъемки – цифровая фотограмметрическая обработка, позволяющая определять геометрические, количественные и другие свойства объектов на поверхности земли по фотографическим изображениям, получаемым с помощью летательных аппаратов любых видов. В настоящее время обработка полученных изображений ведется с помощью компьютерной техники и программ для фотограмметрической обработки изображений. При этом дополнительно выполняются коррекции перспективы, дисторсии и иных оптических искажений, цветокоррекция полученных снимков.

Фотографическое изображение преобразовывается в цифровую форму путем сканирования. При этом изображение делится на определенное количество пикселей равных площадей. Каждый такой площадной объект содержит достаточную информацию в отношении цвета и плотности цвета. В цифровой фотограмметрии точность получения результатов возрастает с увеличением разрешения сканирования. Чем меньше размер пиксела, тем точнее результат.

Широко используемым продуктом цифровой обработки являются ортофотопланы. Ортофотоплан - это фотографический план местности на точной геодезической опоре, полученный путём аэрофотосъёмки с последующим преобразованием аэроснимков (из центральной проекции в ортогональную) на основе эффективного метода их дифференциального ортофототрансформирования.

Космические снимки

Сбор данных для ГИС на базе космических снимков высокого разрешения становится все более популярным благодаря увеличившемуся числу спутников и архивов данных.

Космические съемки выполняют с использованием либо фотографических, либо сканерных систем. Космическая съемка осуществляется специальными приборами - датчиками. Датчики могут быть пассивными и активными, причем пассивные датчики улавливают отраженное или испускаемое естественное излучение, а активные способны сами излучать необходимый сигнал и фиксировать его отражение от объекта. Наряду с этим по-прежнему используются

космические системы с панхроматическим фотооборудованием и многоспектральными фотокамерами, обеспечивающими высокое разрешение и геометрическую точность.

Для дистанционного зондирования Земли используются ультрафиолетовый, видимый или световой диапазон, ближний или фотографический, инфракрасный, тепловой инфракрасный и микроволновый диапазоны волн электромагнитного излучения. Результаты дистанционных измерений, осуществляемых с помощью бортовой информационно-измерительной аппаратуры аэрокосмической системы, представляют собой регистрацию в аналоговой или цифровой форме характеристик электромагнитного излучения, отраженного от участков земной поверхности или собственного излучения этих участков.

Для сбора геопространственных данных больше подходят объединенные данные мультиспектральных и панхроматических снимков или панхроматические снимки. Сегодня геометрическое разрешение панхроматических снимков в 4 раза выше мультиспектральных.

Виды космических материалов очень разнообразны. Особо следует отметить снимки сверхвысокого разрешения, которые пользуются повышенным спросом. С 2001 г. спутник QuickBird, QuickBird-2 (США) поставлял самые высокодетальные (61 см) изображения Земли, доступные коммерческим пользователям. В 2007 году был запущен Спутник WorldView-1, который обеспечивает большую детальность получаемых изображений Земли из космоса. Пространственное разрешение снимков – 50 см в надири и 59 см при угле отклонения в 25 градусов; точность географической привязки 6,5 м. В 2008 г. был запущен коммерческий спутник GeoEye-1 субметрового разрешения. Сенсоры GeoEye-1 совершенствует линейку современных оптических спутниковых систем по ряду параметров: сверхвысокая детальность изображений (разрешение 0,41 м в панхроматическом и 1,65 м в мультиспектральном режимах), высокая точность привязки снимков (3 м при съемке в надири), высокая производительность.

В результате этих запусков не только увеличивается разрешение получаемых данных, возрастает также точность позиционирования (по оценкам специалистов, до 4-5 м) и возможность извлечения качественных и количественных данных (за счет увеличения числа спектральных диапазонов, оптического и пространственного разрешения, уменьшения периодичности съемки). Значительно возрастает объем ежедневно получаемой информации как за счет

увеличения числа аппаратов, так и в результате улучшения их технических параметров (объем бортовой памяти, скорость передачи, ширина полосы съемки).

Для сбора геопространственных данных важны разрешение снимков, радиометрическое качество, высота солнца и угол обзора. При большом угле отклонения от надира невозможно обозревать городские улицы сверху вниз, при низкой высоте солнца появляются тени, осложняющие дешифрирование улиц. Важны даже атмосферные условия. Атмосферная дымка, тени от облаков и дымка вокруг облаков снижают качество снимков. Данные ДЗЗ, полученные с космического аппарата, характеризуются большой степенью зависимости от прозрачности атмосферы. Поэтому на космических аппаратах используется многоканальное оборудование пассивного и активного типов, регистрирующие электромагнитное излучение в различных диапазонах.

Использование данных дистанционного зондирования является наиболее эффективным методом получения пространственной информации об объектах. Данные ДЗЗ объективны, достоверны, наглядны, при этом затраты на получение информации об исследуемой местности существенно ниже стоимости проведения наземных работ.

Базовые наборы геопространственных данных

В контексте построения инфраструктуры геопространственных данных (ИГ ПД) установления единственной среды предусматривает создания соответствующих информационных ресурсов и их использования [36]. Базовые наборы геопространственных данных являются существенной составляющей инфраструктуры геопространственных данных.

Базовые наборы геопространственных данных определяются как такие, которые:

- 1) предназначены для использования в геоинформационной системе разной отрасли человеческой деятельности;
- 2) являются основой создания производных наборов данных, таких как слои или подтипы данных в геоинформационной системе;
- 3) имеют однозначное определение на местности;
- 4) имеют относительно длительное существование во времени;
- 5) имеют прагматичное значение.

Интернет-источники цифровых пространственных данных

Все большее количество объектов реального мира представляется *готовыми моделями* в цифровом формате. Значительная часть этой информации пробивает себе дорогу на рынок информации. Часть этой информации можно приобрести у поставщиков информации за деньги, другую часть информации можно получить в организациях государственного и местного управления бесплатно или за цену.

Интернет является неоценимым источником поиска и получения цифровых данных. В Интернете можно найти множество пространственных и табличных данных, а также снимков. Хорошими адресами для начального поиска данных являются, например, <http://earth.google.com/>, <http://www.geographynetwork.com/data/> или <http://www.scanex.ru/ru/data/>. В сети Интернет доступ к пространственной информации предоставляют картографические и геоинформационные Web-сервисы.

При определении того, стоит ли приобретать цифровые данные, необходимо знать о происхождении, пространственном охвате, надежности данных, масштабе данных, разрешении, точности и формате данных, актуальности и обновляемости, условиях и стоимости получения, приобретения. Для этого следует получить *метаданные* (данные о данных) в форме словаря данных или отчета о качестве данных от их поставщика. Метаданные должны содержать соответствующую общую информацию о данных. Развитое программное обеспечение ГИС содержат инструменты формализованного описания метаданных.

Общегеографические и тематические карты

Общегеографические карты масштаба от 1:200000 до 1:1000000 и мельче карты содержат разнообразные сведения о рельефе, гидрографии, почвенно-растительном покрове, населенных пунктах, хозяйственных объектах, путях сообщения, линиях коммуникаций, границах.

Тематические карты создаются по определенной предметной области. Среди тематических выделяют карты природы, населения, экономики и др. [16].

Карты природы включают карты геологического строения и ресурсов недр, геофизические, рельефа земной поверхности и дна океанов, метеорологические и климатические, гидрологические и океанографические, почвенные, геоботанические, ландшафтные, зоо-

географические, медико-географические, общие физико-географические, охраны природы, экологические карты.

Карты экономики включают карты промышленности, карты сельского хозяйства, карты лесного хозяйства, карты транспорта и другие, а также комплексные атласы различной тематики. Политические, административные и исторические карты выделяют в отдельную группу.

Документация землеустройства

Документация землеустройства - это утвержденные в установленном порядке текстовые и графические материалы, которыми регулируется использование и охрана земель государственной, коммунальной и частной собственности, а также материалы обследования и разведки земли, авторского надзора, за выполнением проектов и тому подобное. Документация землеустройства разрабатывается в виде программы, схемы, проектов, специальной тематической карты, атласов, технической документации.

К технической документации относятся кадастровые планы, индексные карты. Кадастровые планы содержат границы земельных участков (недвижимой собственности), которые не отображаются на топографических планах. Идентификация земельных участков выполняется путем присвоения кадастровых кодов (номеров). Для уникальной идентификации земельных участков создаются индексные карты.

Управление использованием земли осуществляется системой, которая называется зонингом. Основу зонинга составляют зонинговые карты и зонинговые правила.

Градостроительная документация

Градостроительная документация – это утверждены текстовые и графические материалы, которыми регулируется планирование, застройка и другое градостроительное использование территории. Документация обязательно состоит из текстовых и графических материалов. В основу управления градостроительным развитием территории положен иерархический подход, который предусматривает разработку в определенной последовательности градостроительной документации на общегосударственном, региональном и местном уровнях.

Общегосударственный уровень градостроительной документации включает разработку Генеральной схемы планирования территории Украины, главными заданиями которой является определение стратегии территориального развития нашего государства и ее отдельных регионов

Региональный уровень градостроительной документации решает вопрос территориального развития отдельных регионов Украины, использует схемы планирования отдельной части территории Украины, области и административных регионов.

Местный уровень градостроительной документации должен основываться на решениях регионального уровня. На местном уровне создаются генеральные планы конкретных населенных пунктов, схемы планирования территории сельсоветов, детальные планы застройки.

Таблицы и текстовые документы

Источниками семантических данных могут быть таблицы. Информационный продукт не всегда представляет собой карту. Он также может включать список цифр, таблицу или отчет. Таблица или список могут также составлять целый информационный продукт. Все эти элементы входят в категорию так называемых табличных данных. Электронные таблицы, базы данных и таблицы текстовых файлов являются частью широкой категории табличных данных. Каждая таблица должна иметь название, соответствующие заголовки колонок, типичные записи и указание на источник этих данных. Если существуют установленные форматы отчетов, обязательно определите их на данном этапе. Обычно отчеты можно автоматически конвертировать в нужные форматы.

К текстовым документам предметной или тематической информации относятся технические, экономические, статистические, социологические, демографические и другие виды данных; данные государственной статистической отчетности. Текстовая информация более повествовательного характера часто является важной составляющей информационного продукта. Текстовые файлы могут быть в формате Adobe.PDF, Microsoft Word или в формате .txt.

3.1.2 Характеристики данных

Географические данные имеют многие характеристики. Согласно [37] существенными для сбора данных являются масштаб, разрешение, картографическая проекция, допуск ошибки.

Масштаб

Масштаб данных отражает детальность (определенный объем информации) и относительную точность отображения объектов на карте. Чем больше масштаб, тем более точен и подробен набор данных.

В некоторых случаях может потребоваться хранение данных в базе в нескольких масштабах. Если в базе данных представлено больше одного масштаба, это должно быть соответствующим образом документировано в метаданных.

Масштаб влияет как на стоимость, так и на точность конечной базы данных. Например, число листов карты, необходимых для покрытия одной и той же области в базе данных, растет экспоненциально с увеличением масштаба. Поэтому для картирования территории в масштабе 1:6000 требуется в 16 раз больше листов, чем картирование в масштабе 1:24000. С ростом числа листов и детальности, необходимых для определенного приложения, растет и стоимость работ.

Разрешение

Пространственное разрешение определяется как размер самых мелких элементов, поддающихся нанесению на карту или выборке в данном масштабе. Разрешение карты напрямую связано с ее масштабом. По мере уменьшения масштаба карты разрешение уменьшается; границы элементов должны быть сглажены, упрощены или просто не должны быть отображены. Существуют минимальные размеры полигона и длина линии, которые можно отобразить в данном масштабе. Элементы с более низкими разрешениями сливаются с окружающими данными, преобразуются в точку или удаляются.

Данные должны иметь минимально достаточное разрешение, чтобы на их основе можно было создавать информационные продукты. Например, данные о земельных участках в городе должны иметь высокое разрешение, а Web-приложение, отображающее транспортные

маршруты между штатами, должно быть мелкомасштабным. Разрешение также влияет на ошибку данных.

Картографическая проекция

Картографическая проекция преобразует трехмерную поверхность Земли в плоскую двумерную поверхность. Этот процесс создает картографические искажения, относящиеся к расстоянию, площади, форме геометрической фигуры или направлению. В результате, все плоские карты имеют некоторую степень пространственного искажения. Тип используемой проекции определяет степень и тип искажения на карте. Конкретная картографическая проекция может сохранить неискаженным одно свойство за счет других или, в качестве компромисса, отобразить несколько свойств с уменьшенной точностью. Необходимо знать искажение, которое нанесет наименьший ущерб качеству вашей базы данных.

Даты (Datum) - важный аспект, связанный с проекцией. Даты - это набор базовых параметров и контрольных точек, используемых для точного задания трехмерной формы Земли. Даты задают начало отсчета и ориентацию для линий широты и долготы. Даты дают базовую ссылку для определения местоположений на поверхности Земли. Для эффективного использования пространственных данных, полученных с карты, нужно знать проекцию этой карты и даты.

Степень искажения данных, которую вы получите от картографической проекции, связана с масштабом. Чем больше географическая область, покрываемая картой (чем меньше масштаб), тем большее искажение от проекции вы получите.

Допуск ошибки

Допуск ошибки определяет, какой размер ошибки допустим, а какой — недопустим. Поскольку ошибка связана с разрешением и масштабом, она также напрямую привязана к стоимости: уменьшение процента ошибки стоит денег.

Существует четыре типа ошибок.

1. *Ссылочная ошибка* относится к ошибке в определении или привязке ссылок. Например, правильно ли указаны соответствующие названия улиц и номера домов?
2. *Топологическая ошибка* происходит при разрыве в необходимых привязках, например, если полигоны не замкнуты или уличные сети разорваны.

3. *Относительная ошибка* — это ошибка в расположении двух объектов по отношению друг к другу. Например, важно знать местоположение коллектора по отношению к владению и дороге.
4. *Абсолютная ошибка* — это ошибка в определении реального положения какого-либо объекта в мире.

3.1.3 Системный подход к предварительной обработке исходных данных

Множество данных геоинформационных систем собирается с помощью разных технических средств, разных технологий и в разных условиях. Данные отражают различные характеристики и свойства. Они могут иметь различные размерности, разное количество значащих цифр, разную точность и т.д. Для того чтобы разнородные данные и модели можно было обрабатывать совместно, они должны быть предварительно упорядочены и представлены в форме, воспринимаемой единой системой данных [38].

Системы данных – это взаимосвязанные совокупности унифицированных, упорядоченных, организованных данных, описываемых по единым правилам. Это могут быть наборы файлов, данные в базе данных, интегрированная информационная основа ГИС.

Использование системного подхода имеет целью подготовки данных для дальнейшей интеграции их в системы данных, с помощью которых удовлетворяются потребности информационных процессов и решаемых задач. Системный подход к обработке данных требует представления исходных данных как компонентов будущих или уже существующих систем данных.

Каждый вид данных в зависимости от степени или параметров соответствия системным требованиям может требовать выполнения одной или нескольких процедур обработки, или не требовать таковых. Наибольшее применение находят следующие виды процедур предварительной обработки исходных данных: первичная обработка, локализация географических объектов, оцифровка, трансформирование данных, конвертирование данных, унификация, классификация, идентификация, стратификация и другие.

Первичная обработка

При первичной обработке используются общие и специальные средства подготовки данных [39].

Общими могут быть следующие средства подготовки данных.

Фильтрация данных. Загруженные (или предварительно обработанные) данные могут быть избыточными по количеству записей или содержанию отдельных полей с точки зрения задач пользователя. Для удаления избыточных по ряду выделенных признаков записей может быть использована процедура фильтрации данных.

Нормализация данных. Использование отдельных инструментов требует предварительной нормализации данных, то есть приведения интервалов значений отдельных множеств к интервалам с предопределенными границами. Ряд продуктов имеют инструменты, которые реализуют процедуру нормализации.

Удаление противоречий. В анализируемом наборе данных можно выделить входные и выходные поля. Значения входных полей полностью определяют значения выходных полей. Противоречия возникают, если есть группы записей, в которых значения во входных полях полностью совпадают, а в выходных – различаются.

Удаление дубликатов. В данных могут встречаться записи с одинаковыми входными факторами и одинаковыми выходными, т.е. дубликаты. Такие данные несут избыточность. Удаление дубликатов осуществляется с помощью специальной процедуры.

Разделение. Процедура разделения позволяет разделить исходный набор данных на несколько непересекающихся подмножеств. Если необходимо использовать простые принципы разделения данных, процедура разделения эффективнее, чем параллельное применение нескольких процедур.

Интерполяция. Если в столбце некоторые данные отсутствуют в силу каких-либо причин, обосновывается один из двух путей работы с данными: а) убрать из обработки все строки, которые содержат пропущенные данные, б) заполнить пропущенные данные, используя процедуру интерполяции.

Примеры специальных средств подготовки данных:

- первичная обработка данных ДЗЗ включает радиометрическую калибровку, географическую привязку, геометрическую коррекцию изображений и др.

- первичная обработка GPS данных включает сохранение данных с приемника GPS в файл журнала (Log File), фильтрацию входных данных GPS по времени, расстоянию или отклонению.

Локализация географических объектов

Установление пространственных отношений между процессами, явлениями и объектами на поверхности Земли является одной из главных задач геоинформационных систем. Эта проблема решается посредством локализации географических объектов.

Пространственной локализацией географических объектов называют процесс установления пространственных отношений между характеристиками геоданных путем соотнесения разных видов информации к одному из наиболее устойчивых ее видов. Этим видом являются координаты в общеземной или локальной системе координат. По существу, локализация в геоинформатике означает процесс привязки данных к системе координат реального земного пространства.

Оцифровка

Оцифровка (Digitizing) - синоним цифрование, дигитализация, – есть процесс представления объекта, изображения, документа, звука или сигнала (как правило, аналогового сигнала) посредством дискретного набора точек или элементов. Другими словами, это процесс перевода аналоговых данных в цифровую форму, доступную для существования в цифровой компьютерной среде. В результате оцифровки получается так называемое "цифровое изображение" объекта или "цифровая форма" сигнала. Оцифровка текстов используется для преобразования изображения текстов в цифровую форму с помощью оптического распознавания символов.

Оцифровка имеет огромное значение для вычислительного процесса. Дигитализация является основным способом представления изображений в формате, пригодном для передачи и компьютерной обработки.

Оцифровка происходит в два этапа:

- 1) Дискретизация – а) разделение на аналоговые сигналы и на регулярные интервалы времени (частоты дискретизации), б) чтение значений сигнала в точке.

- 2) Квантование - округление до фиксированного набора чисел (например, целых).

В географических информационных системах дигитализация есть инструмент цифрового представления географических объектов в растровой или векторной форме. Дигитализация является инструментом создания электронных карт из различных географических изображений или путем оцифровки традиционных бумажных аналоговых карт. Различают следующие две технологии оцифровки бумажных карт.

- 1) оцифровка с помощью графического устройства - дигитайзера и ручного обвода объектов карты (tablet-based digitizing);
- 2) оцифровка путем сканирования оригиналов для получения растровых копий и последующей векторизации растровой картографической подложки по дисплею манипулятором "мышь" (on-screen digitizing).

Векторизацией называется также процесс конвертирования растровой графики в векторную графику с использованием программного и компьютерного обеспечения.

По степени автоматизации различают ручную (manual), полуавтоматическую (semi-automated) и автоматическую (automatic) дигитализацию. Процессы дигитализации обслуживаются программными средствами, которые называются графическими редакторами.

Трансформация данных

Трансформация данных (Data transformation) - перенесение данных с одного носителя данных на другой без существенного изменения содержащейся в них информации. Трансформация данных позволяет по заданным наперед правилам изменения структуры данных, строить из одних экземпляров объектов данных другие экземпляры. Чаще всего трансформация используется для предварительной обработки информации без построения конкретной модели.

В геоинформационных системах часто возникает необходимость решения задач:

- преобразования данных между координатными системами (преобразование сдвигом, преобразование сдвигом и поворотом, аффинное преобразование, полиномиальные преобразования, преобразование "резиновый лист")

- преобразования данных между различными форматами данных (формат данных есть внутренняя структура файла, которая позволяет использовать компьютер).

Организации во всем мире производят пространственные данные в различных форматах данных. При сборе данных эти форматы бывают двух разновидностей:

- родные форматы программного обеспечения ГИС,
- неродные (внешние) форматы программного обеспечения ГИС.

Например, для ArcGIS:

- родные форматы, такие как шейпфайлы (shapefiles), покрытия (coverages) или какие-либо элементы базы геоданных (geodatabase), такие как класс пространственных объектов (feature class) или таблицы (table);
- неродные или внешние форматы, такие как DWG/DXF, DLG, MicroStation Design, MapInfo MID/MIF, MapInfo TAB и др.

При необходимости использовать данные, хранящиеся во внешних форматах, следует использовать расширение функциональной совместимости данных, например, расширение Data Interoperability в ArcGIS. Это расширение позволяет читать и записывать данные в более чем 70 различных форматах.

Конвертирование данных

Конвертирование данных (Data conversion) - изменение формы представления данных в соответствии с определенными правилами при сохранении содержащейся в них информации. Например, полноцветные изображения могут быть преобразованы в градациях серого, преобразования текстовых файлов из одной кодировки символов в другую, конверсия метрических единиц.

В ArcGIS набор инструментов Conversion toolbox содержит инструменты, которые могут конвертировать данные между различными формами. Выбор метода будет зависеть от формы и качества существующих данных, формата данных из внешних источников и установленных стандартов.

Унификация

Разнообразие технологий и методов сбора данных порождает разнообразие типов данных, которые впоследствии необходимо обраба-

тывать. Обработать множество различных данных неудобно и неэффективно. Для упрощения процесса обработки, хранения и обмена разнородные данные приводят к структурному единому виду, который используется при последующей обработке информации. Такие данные называют унифицированными [38]. Процедура сведения разнородных видов и структур данных к единому виду и структуре называется унификацией. В ходе унификации данных осуществляется построение единой структуры информационной модели.

Классификация

В процессе организации данных осуществляется их классификация. Она позволяет соотносить различные модели и их характеристики к разным классам, подклассам и типам, что дает возможность систематизировать исходные наборы данных и использовать свойства классов при анализе конкретных геоинформационных данных. Это одна из важнейших функций организации данных, геоданных. Таким образом, организованные данные — это классифицированные данные.

Идентификация

Идентификация - это процедура присвоения наборам данных специальных кодов - идентификаторов. Эта процедура обеспечивает взаимно однозначное соответствие между данными и идентификаторами, исключает повторение одинаковых идентификаторов, то есть обеспечивает непротиворечивость моделей и их характеристик.

Стратификация

В ГИС один из общепринятых принципов организации пространственной информации называют послойным принципом. Суть его состоит в том, что данные о территории организуются в виде набора тематических слоев, отвечающих конкретным потребностям и содержащих небольшое число типов однородных данных, объединяемых общей тематикой. Например, слой почв, слой гидрографической сети, слой транспортной сети, слой зданий, слой инженерных сетей и так далее.

Стратификация - это создание слоев, которые могут иметь несколько значений в разных областях применения. Стратификация данных в ГИС - это процедура, которая разбивает систему данных на

страты или слои, удобные для обработки и анализа. Стратификация - это не просто организация данных, а создание инструмента анализа и обобщения данных на разных территориальных уровнях. Это процесс задания некой структуры для эффективной работы геоинформационной системы [38].

Пространственные объекты характерны тем, что имеют графическую форму представления. Графическая форма представления пространственных объектов должна решать две задачи: показывать взаимное расположение объектов и связи между ними (топология); давать возможность количественной оценки геометрических характеристик объектов и их положения в выбранной системе координат (метрика). Графическая форма представления пространственных объектов является в то же время основой для стратификации данных.

* * *

Реализация изложенных и других процедур создает основу для последующей системной организации и анализа пространственных данных.

3.1.4 Контрольные вопросы и задания для самостоятельной работы

- 1) Охарактеризуйте аналоговые и цифровые топографические карты как источник географических данных.
- 2) Охарактеризуйте данные топографических съемок, измерений электронными тахеометрами и приемниками глобальной системы позиционирования как источник географических данных.
- 3) Охарактеризуйте материалы дистанционного зондирования Земли как источник географических данных.
- 4) Приведите основные характеристики географических данных.
- 5) В чем заключается системный подход к предварительной обработке исходных данных?
- 6) В чем заключается оцифровка аналоговых топографических карт?
- 7) Что понимается под стратификацией географических данных?

Раздел 3.2

КЛАССИФИКАЦИЯ

3.2.1 Общие положения классификаций

Для представления общих положений классификаций использованы концепции, изложенные в общепризнанном руководстве Гради Буча [9].

3.2.1.1 Классификация — средство упорядочения знаний

Классификация – логическая операция, состоящая в разделении всего изучаемого множества предметов по обнаруженным сходствам на отдельные группы или подчиненные множества, называемые классами. При классификации выполняется распределение предметов какого-либо рода на взаимосвязанные классы согласно наиболее существенным признакам, присущим предметам данного рода и отличающим их от предметов другого рода, при этом каждый класс занимает в получившейся системе определенное постоянное место и, в свою очередь, делится на подклассы.

Целью классификации является нахождение общих свойств объектов. Классифицируя, мы объединяем в одну группу объекты, имеющие одинаковое строение или одинаковое поведение. Правильно составленная классификация отображает закономерности развития классифицируемых объектов, глубоко вскрывает связи между изучаемыми объектами и помогает исследователю ориентироваться в самых сложных ситуациях, служит основой для обобщающих выводов и прогнозов.

Классификация нужна в любой области знаний, в любой целесообразной деятельности человека. Классификация – это один из фундаментальных процессов в науке. Факты и явления должны быть упорядочены, прежде чем мы сможем их понять и разработать общие принципы, объясняющие их появление и видимый порядок. Разумная классификация – часть любой точной науки. Построение содержательной классификации наблюдаемых объектов и ситуаций является неотъемлемой задачей науки. Такая классификация существенно облегчает понимание основной проблемы и дальнейшее развитие научной теории.

Успешной классификации обычно сопутствуют открытия и изобретения. С помощью открытий мы распознаем ключевые понятия и механизмы. С помощью изобретений мы конструируем обобщенные понятия, а также новые механизмы, которые определяют правила взаимодействия объектов. Поэтому открытия и изобретения являются неотъемлемой частью успешной классификации.

К сожалению, пока не разработаны строгие методы классификации и нет правил, позволяющих выделять классы и объекты. Нет таких понятий, как "совершенная структура классов", "правильный выбор объектов". Как и во многих технических дисциплинах, выбор классов является компромиссным решением. Тем не менее, имеется богатый опыт классификации в разных областях науки, на основе которого разработаны методики объектно-ориентированного анализа и прикладного анализа. Каждая такая методика предлагает свои правила идентификации классов и объектов. Ниже приводятся некоторые примеры известных классификаций.

Пример классификации в химии.

Менделеев Д.И. в 1869 году установил периодический закон, который определял зависимость различных свойств химических элементов от заряда атомного ядра и мог предсказывать свойства еще не открытых элементов. Графическим выражением периодического закона стала классификация химических элементов в форме "Таблицы Менделеева". Появление периодической системы Менделеева Д.И. открыло новую, подлинно научную эру в истории химии и ряде смежных наук. Взамен разрозненных сведений об элементах и соединениях появилась стройная система, на основе которой стало возможным обобщать, делать выводы, предвидеть.

Пример классификации в биологии.

В современной биологии термин "классификация" обозначает установление иерархической системы категорий на основе некоторых предопределенных связей между организмами. Наиболее общее понятие в биологической таксонометрии – мир. В порядке убывания общности иерархическая система категорий представляется следующим образом: мир, тип, класс, подтип, сорт, семейство, род, вид. В современной классификации живых организмов определяются группы организмов, имеющих общее генетическое наследство.

3.2.1.2 Сложность классификации

Определение классов и объектов - одна из самых сложных задач объектно-ориентированного проектирования. Сложность классификации объясняется двумя важными причинами.

Во-первых, нет определения "совершенная классификация", хотя естественно некоторые лучше других. Существует столько способов деления мира на объектные системы, сколько ученых принимаются за эту задачу: способ классификации определяется целью, к которой мы стремимся.

Во-вторых, разумная классификация требует большой творческой энергии и проницательности. Существенные признаки, соответствующие некоторым закономерностям определенного класса, не лежат на поверхности, а должны быть найдены посредством анализа. Классификация предполагает сложный системный анализ, заставляет искать содержательные признаки. Этим подчеркивается, что разумная классификация – работа интеллектуальная.

Лучший способ ведения классификации – последовательный итеративный процесс. Последовательный итеративный подход непосредственно определяет процедуру конструирования иерархии классов и объектов при разработке сложного программного обеспечения. На практике, обычно за основу берется какая-либо определенная структура классов, которую постепенно совершенствуют. И только на поздней стадии разработки, когда уже получен некоторый опыт использования такой структуры, можно критически оценить качество получившейся классификации. Основываясь на полученном опыте, можно создать новые подклассы из уже существующих или слить несколько существующих в один. Возможно, в процессе разработки были найдены новые общие свойства, ранее не замеченные, и можно определить новые классы.

3.2.1.3 Подходы классификации

Исторически сложилось три подхода классификации:

- 1) классическое распределение по категориям (группирование по свойствам),
- 2) концептуальное объединение (группирование по некоторой концепции),
- 3) теория прототипов (группирование объектов по некоторым признакам схожести с прототипом).

Классическое распределение по категориям

В классическом подходе распределение по категориям выполняют в соответствии с принципом: "все вещи, обладающие данным свойством или совокупностью свойств, формируют некоторую категорию. Причем, наличие этих свойств является необходимым и достаточным условием, определяющим категорию".

Для метода классического распределения по категориям наличие *свойства* является основным критерием схожести объектов. При этом объекты можно разделить на непересекающиеся множества, определив для каждого, обладает ли он конкретным свойством или нет.

Конкретные свойства, которые необходимо выделить при классификации определяются решаемой проблемой. Например, цвет автомобиля является важным свойством для какой-нибудь заводской системы контроля, но не имеет значения для программной системы, управляющей городскими светофорами. Вот почему можно утверждать, что нет абсолютных критериев совершенства, хотя для конкретного примера некоторые структуры более пригодны, чем остальные.

Концептуальное объединение

Концептуальное объединение — более современная вариация классического подхода. Она возникла из попыток формального представления знаний. При таком подходе сначала формируют концептуальные описания классов, и затем объекты классифицируются согласно описанию.

Теория прототипирования

Есть ситуации, в которых не работают методы классического распределения по категориям и концептуальное объединение. Существуют некоторые абстракции, которые не имеют ни четких свойств, ни четкого определения. Существуют категории, которые не соответствуют классическим образцам, так как нет признаков, свойственных всем составляющим. При такой классификации класс определяется одним объектом-прототипом. Объект можно включить в класс при условии, что он определенным образом похож на прото тип.

3.2.1.4 Значение классификации для геоинформационных систем

Отсутствие классификаций объектов и свойств является непреодолимым препятствием в создании эффективных информационных систем. Особенно это относится к *геоинформационным системам*, которые оперируют с огромными совокупностями и большим разнообразием объектов реального мира, их свойствами, связями и отношениями. Классификация лежит в основе эффективной работы ГИС.

В объектно-ориентированном анализе определение общих свойств объектов помогает найти общие ключевые абстракции и механизмы, что, в свою очередь, приводит к более простому проекту системы. Классификация затрагивает многие аспекты объектно-ориентированного проектирования. Она помогает определить обобщенную, специализированную и собирательную иерархию классов. Определение общих форм взаимодействия объектов позволяет найти механизм, который может стать стержнем реализации проекта и анализа. Она помогает правильно определять модульную структуру. Расположение объектов в одном или разных модулях зависит от степени общности объектов.

Основным средством обеспечения непротиворечивого описания данных является стандартизация и унификация средств и методов их представления. Средством стандартизации и унификации данных является классификация объектов, свойств и отношений. Стандартизация, унификация и классификация объектов, свойств и отношений лежат в основе информационной совместимости разных геоинформационных систем.

3.2.2. Базовые понятия единой системы классификации технико-экономической информации

Основу компьютерной информационной системы составляет информационная база, которая является совокупностью упорядоченной информации, используемой при функционировании информационной системы. Информационная база имеет целью обеспечить взаимообмен информацией между структурными единицами компьютерной информационной системы, а также информационной системой разных уровней управления.

Успех создания единой информационной базы в значительной мере определяется унификацией и стандартизацией ее составляющей. Здесь классификации и кодировке технико-экономической информации отводится особая роль, поскольку они являются средствами, которые обеспечивают взаимный обмен информацией между человеком и компьютером. Классификация и кодировка — это две неотделимые части одного процесса — перевода разнообразной информации из естественного языка на формализованный язык компьютерной информационной системы.

Государственный стандарт 17369-85 устанавливает основные термины и определения понятий, применяемых в единой системе классификации и кодирования технико-экономической информации (ЕСКК ТЭИ) [40].

3.2.2.1 Термины и определения классификации

В ГОСТ 17369-85 определены следующие термины и определения классификации.

- Классификация - разделение множества объектов на подмножества по их сходству или различию в соответствии с принятыми методами.
- Система классификации - совокупность методов и правил классификации, а также ее результат.
- Объект классификации - элемент классифицируемого множества.
- Признак классификации - свойство или характеристика объекта, по которому производится классификация.
- Значение признака классификации - качественное или количественное выражение признака классификации.
- Классификационная группировка - подмножество объектов, полученное в результате классификации.
- Иерархический метод классификации - последовательное разделение множества объектов на подчиненные классификационные группировки.
- Фасетный метод классификации - параллельное разделение множества объектов на независимые классификационные группировки
- Степень классификации - этап классификации при иерархическом методе, в результате которого получается совокупность классификационных группировок.
- Глубина классификации - число ступеней классификации.

3.2.2.2 Термины и определения кодирования

В ГОСТ 17369-85 определены следующие термины и определения кодирования.

- Код - знак или совокупность знаков, принятых для обозначения классификационной группировки и (или) объекта классификации.
- Кодирование - образование и присвоение кода классификационной группировки и (или) объекту классификации.
- Система кодирования - совокупность методов и правил кодирования классификационных группировок и объектов классификации заданного множества.
- Алфавит кода - система знаков, принятых для образования кода.
- Основание кода - число знаков в алфавите кода.
- Цифровой алфавит кода - алфавит кода, знаками которого являются цифры.
- Буквенный алфавит кода - алфавит кода, знаками которого являются буквы алфавитов естественных языков.
- Буквенно-цифровой алфавит кода - алфавит кода, знаками которого являются буквы алфавитов естественных языков и цифры.
- Разряд кода - позиция знака в коде.
- Длина кода - число знаков в коде без учета пробелов.
- Структура кода - условное обозначение состава и последовательности расположения знаков в коде.
- Последовательный метод кодирования - образование кода классификационной группировки и (или) объекта классификации с использованием кодов последовательно расположенных подчиненных группировок, полученных при иерархическом методе классификации, и его присвоение.
- Параллельный метод кодирования - образование кода классификационной группировки и (или) объекта классификации с использованием кодов независимых группировок, полученных при фасетном методе классификации, и его присвоение.
- Порядковый метод кодирования - образование кода из чисел натурального ряда и его присвоение.
- Серийно-порядковый метод кодирования - образование кода из чисел натурального ряда, закрепление отдельных серий или диапазонов этих чисел за объектами классификации с одинаковыми признаками и его присвоение.

- Перекодирование - присвоение закодированной классификационной группировке или закодированному объекту классификации нового кода.
- Перекодировочная таблица - таблица соответствия кодов одних и тех же классификационных группировок и (или) объектов классификации из разных классификаторов.

3.2.2.3 Организационные аспекты классификации

В ГОСТ 17369-85 определены следующие организационные термины и определения.

- Классификатор - официальный документ, представляющий систематизированный свод наименований и кодов классификационных группировок и (или) объектов классификации.
- Позиция классификатора - наибольшее количество позиций, которое может содержать классификатор.
- Емкость классификатора - количество свободных позиций в классификаторе.
- Внедрение классификатора - проведение комплекса мероприятий, обеспечивающих применение классификатора в предназначенной сфере деятельности.
- Ведение классификатора - поддержание классификатора в достоверном состоянии.
- Категория классификатора - признак, указывающий на принадлежность классификатора к определенной группе в зависимости от уровня его утверждения и сферы применения.
- Государственный классификатор - классификатор, утвержденный для применения в автоматизированных системах управления различных министерств (ведомств).
- Отраслевой (ведомственный) классификатор - классификатор, введенный в установленном порядке для применения в автоматизированной системе управления отрасли министерства (ведомства).
- Классификатор предприятия - классификатор, введенный в установленном порядке для применения в автоматизированной системе управления предприятия (АСУП).
- Единая система классификации и кодирования технико-экономической информации (ЕСКК ТЭИ) - совокупность взаимосвязанных государственных классификаторов технико-

экономической информации, системы ведения и руководящих нормативных документов по их разработке, внедрению, ведению, совершенствованию и контролю за внедрением.

- Регистрация классификатора - присвоение утвержденному классификатору регистрационного номера и запись необходимых сведений о нем в реестр.

3.2.2.4 Особенности создания системы классификации

Создание системы классификации на основании базовых понятий имеет некоторые особенности [41].

Система классификации определяется и характеризуется использованным методом классификации, признаками классификации (положенным в основу выделения классификационной группировки), их последовательностью и количеством уровней (ступеней) классификации, а также количеством группировок (емкостью). Метод классификации — это в сущности совокупность правил создания системы классификационной группировки и их связки между собой.

Иерархический метод классификации

Иерархический метод классификации характеризуется тем, что начальное множество объектов последовательно разделяется на группировку (классы) первого уровня разделения, дальше — на группировку следующего уровня и так далее. Каждая группировка разделяется на группировку следующего уровня. Совокупность группировки образует при этом иерархическую древовидную структуру.

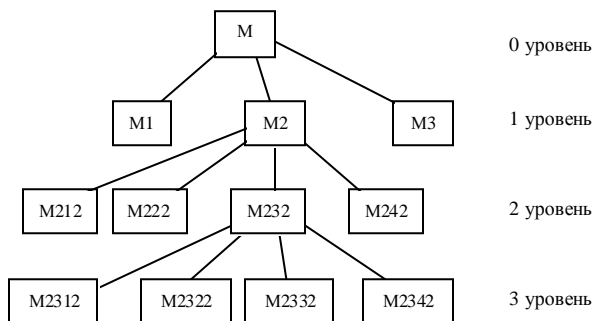


Рис.3.2.1 - Схема группировки иерархического метода классификации

Количество уровней (ступеней) классификации определяет глубину классификации. Каждая группировка может разделяться на группировку низшего уровня с использованием своего признака; глубина классификации в каждой ветви иерархической структуры может быть разной.

При использовании иерархического метода необходимо соблюдать такие основные правила:

- 1) разделение каждой группировки выполняется лишь по одному основному признаку;
- 2) пересечение множеств объектов не допускается; на каждом уровне классификации группировки не должны повторяться;
- 3) классификация должна охватывать все множество объектов.

Наиболее существенными и сложными вопросами, которые появляются при использовании иерархического метода классификации, является выбор системы признаков, которые станут основой разделения, а также их последовательность. Выбранные признаки должны являться определяющими в развязывании конкретной технико-экономической задачи, для которой создается эта система классификации.

Позитивной чертой иерархического метода классификации является его приспособленность к ручной обработке, структурность, большая информативность кодов, которые несут смысловую нагрузку. В тоже время, иерархический метод классификации имеет недостатки, которые иногда усложняют его использование:

- метод имеет жесткую структуру, которая предопределена фиксированностью признаков и их последовательностью; изменение хотя бы одного признака приводит к перераспределению классификационных группировок;
- метод не дает возможность агрегировать объекты по любым раньше не предусмотренным произвольным сочетанием признаков, а также в некоторых случаях усложняет автоматизированную обработку, поскольку образуется нестандартное распределение последовательности признаков.

Фасетный метод классификации

При использовании фасетного метода начальное множество объектов может независимо разделяться на классификационную группировку с использованием одного из избранных признаков. Каждый признак фасетной классификации отвечает фасете, который является списком значений именуемого признака классификации. Например, признак "цвет" содержит такой список фиксированных

значений: красный, желтый, зеленый, голубой, белый, черный. Следуют обратить внимание, что в приведенной фасете "цвет" отсутствует цвет оранжевый потому, что он не был объектом классификации.

Система классификации фасетным методом может быть представлена перечнем независимых фасетов (списков), которые содержат значение признаков классификации. Любая группировка в системе классификации может быть определена одним признаком или набором признаков объектов классификации.

При применении фасетного метода классификации следуют соблюдать такие основные правила:

- признаки, которые используются в разных фасетах, не должны повторяться (принцип взаимного исключения фасетов);
- из всяких признаков, которые характеризуют множество объектов классификации, отбираются и фиксируются лишь существенные, которые обеспечивают решение конкретной задачи.

Фасетный метод классификации не имеет недостатков иерархического метода. Он особенно эффективен при функционировании компьютерных информационных систем.

Кодирование

Кодирование предназначено для формализованного описания объектов классификации, чаще всего в виде цифровых кодов. Такое представление приемлемее всего для повышения эффективности автоматизированной обработки экономической информации.

В процессе кодировки экономической информации необходимо решить три основные задачи:

- 1) однозначного обозначения (идентификации) каждого объекта заданного множественного числа,
- 2) кодирование некоторой совокупности свойств (атрибутов) объекта,
- 3) обеспечение эффективности и информационной надежности переработки информации.

Основным требованием, которое ставят к кодировке, является однозначное представление каждого объекта кодируемого множественного числа, то есть каждому объекту множественного числа должен отвечать единственный код.

Код любого объекта может состоять из идентификационной части, информационного блока, который содержит набор кодов, отвечающих свойствам данного объекта, и дополнительных разрядов или блоков, которые обеспечивают защиту всего кода от возможной ошибки.

При кодировке совокупности свойств объектов, то есть при создании информационного блока, рекомендуют применять два основных метода создания кода: последовательной кодировки на основе использования иерархической классификации и параллельной кодировки на основе фасетной классификации.

В каждой ветви иерархической классификации, как правило, используется своя совокупность свойств. В случае последовательного кодирования код группировки создается на основе кода группировки предыдущего уровня добавлением к нему еще одного разряда (или группы разрядов). При использовании последовательного метода логично строится код (кодовая комбинация), который имеет большую информативность. Но код при этом очень громоздок и имеет сложную структуру. Его целесообразнее использовать лишь в тех случаях, когда информация изменяется в незначительных размерах или вовсе не изменяется в течение длительного времени использования классификаторов, например при разработке общегосударственных классификаторов.

Табл. 3.2.1 - Последовательный метод кодирования

Код	XXX	XX	XXXX	X
Группировка	Класс	Подкласс	Сорт	Семейство

При использовании фасетной классификации каждая группировка системы классификации отвечает некоторой совокупности значений свойств объектов. При этом каждая группировка первого уровня разделения отвечает одному значению, второго — значению двух свойств и т.п.

Табл. 3.2.2 - Параллельный метод кодирования

Код	XXX	XX	XXXX	X
Фасета	Фасета 1	Фасета 2	Фасета 3	Фасета 4

Самым простым и самым распространенным методом кодировки объектов первичного множества является порядковый метод. При

использовании этого метода каждый объект классифицированного множества кодируется с помощью текущего номера. Порядковый метод применяется при кодировке однопризнаковых, постоянной и малозначашей номенклатуры. Порядковый метод кодировки очень прост для идентификации. Но существенным недостатком данного метода является отсутствие в коде любой информации об объекте и относительная сложность автоматической обработки информации при получении итогов по группам объектов.

Табл. 3.2.3 - Порядковый метод кодирования

Код	XXXX
Порядковые числа	1,2,3,4,...,9999

Серийно порядковый метод кодирования характеризуется тем, что первичное множество делится на несколько частей (согласно некоторому признаку), и для кодировки объектов каждой части предназначается серия номеров (кодов). Объекты кодируются порядковым номером в пределах отведенной для них серии. Этот метод кодировки используется для объектов, какие имеют два признака.

Табл. 3.2.4 - Серийно-порядковый метод кодирования

Код	XXX	XX	XXXX
Серии порядковых чисел	Серия 1, 1,2,3,...,999	Серия 2, 1,2,3,...,99	Серия 3, 301,302,...,9999

При выборе метода кодировки следует помнить, что этот метод должен обеспечивать:

- в пределах заданного множества однозначное определение объекта;
- необходимую информацию об объекте;

- использования как алфавита кода десятичных цифр и букв азбуки, что удобно для машинной обработки и обработки человеком;
- по возможности наименьшую длину кода, которая упрощает заполнение документов, упрощает их проверку, уменьшает количество ошибки, размеры машинной памяти и время обработки;
- достаточный резерв незанятых кодов, чтобы можно было кодировать новые объекты и группировки, не в нарушение структуры классификатора;
- возможность автоматического контроля ошибки, например, внесением для кода контрольного разряда.

Классификация и кодирование — это неотделимый элемент создания и функционирования компьютерной информационной системы. Целью классификации и кодирования является упорядочение и согласование разных предметов, понятий, свойств или других элементов информации. Использование кодов можно значительно сократить объемы информации и трудоемкость ее обработки на всех этапах технологического процесса автоматизированной обработки данных. Поэтому вопросом разработки рациональных классификаторов во время проектирования информационной системы уделяется большое внимание.

Разработка классификаторов базируется на учете следующих принципов.

- 1) Обеспечение в классификаторе выделения государственного, ведомственного и местного аспектов управления.
- 2) Содержание названий единицы информации должно отвечать требованиям государственных стандартов или руководящих технических документов.
- 3) Классификатор должен обеспечить решение всей задачи в информационной системе при минимальной значительности кодов. Большая значительность кодов усложняет заполнение документов, утрудняет их проверку, увеличивает расходы времени на подготовку информационных массивов.
- 4) Построение системы классификации и кодирования нуждается в поддержке принципа взаимно однозначного соответствия классификационной номенклатуре. Каждый код должен обозначать лишь один объект классификации.
- 5) Система классификации должна обеспечить все виды группирований кодируемых элементов, необходимых для соответствующих расчетов.

- б) Система классификации и кодирования должна обеспечить необходимый резерв с целью внесения номенклатуры, которые только что возникли, без изменения структуры классификатора.
- 7) Важным требованием является стабильность кодов. Кодов объектов, которые выбыли, не должно присваиваться новой позиции
- 8) Система кодирования должна обеспечить возможность обнаруживать ошибку, которая возникает при вводе или записи кодов, программным способом. Этого можно достичь внесением в код контрольного разряда.
- 9) Для обеспечения совместимости информационной системы разных уровней управления коды по ряду номенклатур должны быть единственными для всех объектов управления. Единство кодов для разных уровней управления обеспечивается внедрением Единой системы классификации и кодирования технико-экономической информации.

Требования, которые относятся к построению классификаторов, разнообразны. Учитывая это при разработке классификаторов в каждом конкретном случае необходимо выбирать оптимальный вариант, который позволяет с помощью вычислительной техники переделывать информацию с минимальными расходами.

3.2.2.5 Единая система классификации и кодирования технико-экономической информации

Единая система классификации и кодирования технико-экономической информации (ЕСКК ТЕИ) — это комплекс взаимосвязанных классификаторов технико-экономической информации, приспособленных к обработке средствами вычислительной техники с автоматизированной системой ведения этих классификаторов.

ЕСКК состоит из совокупности взаимосвязанных классификаторов технико-экономической информации, системы их ведения, научно-методических и нормативно-технических документов разработки, ведения и внедрения, а также организации и службы, которая выполняет работу из классификации и кодирования. ЕСКК устанавливает состав и содержание работы из классификации и кодирования технико-экономической информации, единственный порядок планирования и выполнения этой работы в стране.

Основным заданием создания ЕСКК ТЕИ является упорядочение, систематизация, классификация и кодировка разных видов ТЕИ. Перечень объектов ТЕИ, которые отвечают классификационному множеству классификаторов и определяют их виды, охватывает продукцию, которая выпускается в стране, структурную и административную единицу народного хозяйства (отрасль, министерства, ведомства, объединения, предприятия, учреждения), административно-территориальные единицы, трудовые ресурсы и виды деятельности, природные ресурсы, документацию и т.п.

Единая система классификации и кодирования (ЕСКК) обеспечивает поддержку работы программных модулей информационных системы, создает условия для эффективной автоматизированной обработки данных за счет:

- 1) обеспечения информационной совместимости информационных систем обработки данных, которые взаимодействуют в разной отрасли и на разных уровнях управления хозяйством, и повышения эффективности их функционирования.
- 2) обеспечения упорядоченности, систематизации и унификации ТЕИ, и ее формализованного описания с помощью кодов классификаторов, установления однозначности и непротиворечивости названий объектов ТЕИ.
- 3) обеспечения единого понятийного пространства информационной системы за счет средств межуровневого взаимодействия словарей и классификаторов в рамках иерархической и распределенной базы данных;
- 4) проверки семантической целостности данных;
- 5) развития работы из стандартизации, унификации, агрегирования, устранения дублирования разработок;
- 6) организации и проведения работы из международного сотрудничества в отрасли классификации и кодирования.

Единая система классификации и кодирования допускает использование классификаторов таких видов:

- 1) общегосударственные (государственные) классификаторы;
- 2) межотраслевые классификаторы;
- 3) отраслевые классификаторы;
- 4) классификаторы предприятий (локальные).

ЕСКК Украины предусматривает использование и ведение государственных классификаторов:

- СПАТО - Система позначень автономій, територій, областей;
- КОАТУУ - Державний класифікатор об'єктів адміністративно-територіального устрою України;

- КВЕД - Класифікація видів економічної діяльності;
- УКФВ - Український класифікатор форм власності;
- КОПФГ - Класифікатор організаційно-правових форм господарювання;
- ЄДРПОУ - Єдиний державний реєстр підприємств, організацій України;
- ЗКГНГ - Загальний класифікатор галузей народного господарства;
- СПОДУ - Система позначень органів державного управління;
- КОВ - Класифікатор одиниць виміру;
- УСГК - Українська стандартна галузева класифікація;
- УКОЗ - Український класифікатор основних засобів;
- УКУД - Український класифікатор управлінських документів;
- УКП - Український класифікатор продукції;
- других государственных классификаторов.

Общие правила и порядок разработки национальных классификаторов устанавливает Закон Украины "Про стандартизацію" від 1 грудня 2005 року N 3164-IV, ДСТУ 1.2:2003 "Національна стандартизація. Порядок розроблення національних нормативних документів", ДСТУ ISO/IES GUIDE 59:2000 "Кодекс усталених правил стандартизації".

3.2.3 Классификации в геоинформационных системах

3.2.3.1 Объекты классификации в геоинформационных системах

Объектами исследования в геоинформатике являются объекты, явления и процессы реального мира, их свойства и отношения, представляющие их модели. Это могут быть экономические, социальные, экологические, физические и другие явления. Объекты исследования в геоинформатике различаются по пространственному масштабу, по периоду времени исследования и анализа, по характеру существования и функционирования. При этом между этими разнообразными объектами существуют пространственные, временные, функциональные связи, которые также исследуются методами геоинформатики.

Комплексный характер объектов в геоинформатике и разнообразие областей их применения требуют их классификации и разработки различных классификаторов. Для упорядочения результатов исследования в геоинформатике необходима система классификаторов и согласованная методика классификации, которая даст возможность модернизировать существующие и создавать новые классификаторы, исключая противоречие между ними.

Вследствие большого разнообразия объектов классификации в работе [38] выделяют пять подходов при решении классификационных задач, которые определяют пять типов решений.

- 1) К задачам классификации первого типа относятся те, в которых исходные измерения требуется разделить на устойчивые группы. Эти задачи называют задачами классификации без учителя, кластеризации, таксономии, типизации. Этот тип классификаций основан в основном на опытных данных. В ГИС первая задача классификации возникает и решается при первичной обработке информации и при организации информационной основы.
- 2) Задачи классификации второго типа характеризуются тем, что исходные данные уже сгруппированы и требуется оценить их информативность (значимость) относительно совокупности известных эталонов. Такого рода задачи встречаются при распознавании образов, при дешифрировании снимков и т.д. Вторая задача классификации возникает при сборе первичных данных и при использовании ГИС для экспертных решений или оценок.
- 3) К задачам классификации третьего типа относят задачи разбиения. В них исходные измерения или их функции требуется разбить на устойчивые группы в зависимости от их величин. Третья задача классификации также возникает в приложениях ГИС для решения разных задач, например, экологии, землепользования, статистики и т.п.
- 4) Четвертый тип задач классификации называется процедурным. В нем выполняется анализ и классификация процедур преобразования, а не данных. Четвертый тип классификационных задач требует решения при организации и выборе основных технологических процессов ГИС.
- 5) Пятый тип задач классификации называется структурно-аналитическим. С помощью этих задач производится анализ и классификация аналитически - логических связей в системах, алгоритмах и программах. Основу такой классификации составляют методы структурного и системного анализа. Пятый

тип классификационных задач требует решения при организации ГИС как системы, представления (формализации) ее технологических процессов и оптимизации этих технологических процессов.

3.2.3.2 Примеры классификаций

В Украине используются разнообразные классификаторы, которые входят в состав нормативных документов государственной системы классификации и кодировки технико-экономической и социальной информации в Украине. Очертить основную область применения государственных классификаторов почти невозможно, так как они созданы для использования в самой разнообразной отрасли человеческой деятельности.

Дальше приведено общее описание некоторых классификаторов, которые часто используются в геоинформационных системах.

Державний класифікатор об'єктів адміністративно-територіального устрою України

"Державний класифікатор об'єктів адміністративно-територіального устрою України" (КОАТУУ) 1998 г. является составной частью единой системы классификации и кодирования технико-экономической и социальной информации. В соответствии с КОАТУУ все множественное число объектов классификации распределено по территориальному признаку и административной подчиненности на четыре уровня иерархической классификации. К каждому уровню классификации входят объекты, подчиненные объектам предыдущего уровня. Общее кодовое обозначение объектов классификации КОАТУУ имеет десять разрядов.

- 1) *Первый уровень классификации (разряды 1, 2) включает:*
 - АРК;
 - области;
 - города, которые имеют специальный статус, который определяется законами Украины.
- 2) *Второй уровень классификации (разряды 3 - 5) включает:*
 - города областного подчинения;
 - районы Автономной Республики Крым, области;
 - районы в городах, которые имеют специальный статус, который определяется законами Украины.

- 3) *Третий уровень классификации (разряды 6 - 8) включает:*
- города районного подчинения;
 - районы в городах областного подчинения;
 - поселки городского типа; сельские советы.
- 4) *Четвертый уровень классификации (разряды 9, 10) включает:*
- села;
 - поселки.

Идентификационный код каждого объекта строится с использованием серийно порядкового и последовательного методов кодирования. Для объектов первого уровня классификации выделены разряды 1,2, второго уровня — разряды 3-5, третьего уровня — разряды 6-8, четвертого уровня — разряды 9,10. В случае кодирования объектов второго и третьего уровня классификации применяется последовательный метод кодирования. При этом разряды 3 и 6 отведено под признаки, которые указывают на уровень классификации и вид объекта классификации, который кодируется.

- *Признак 2-го уровня (3-й разряд) имеет такое значение:*
 - 1 — города областного подчинения;
 - 2 — районы Автономной Республики Крым, области;
 - 3 — районы городов, которые имеют специальный статус, который определяется законами Украины.
- *Признак 3-го уровня (6-й разряд) имеет такое значение:*
 - 1 — города районного подчинения;
 - 2 — не используется;
 - 3 — районы в городах областного подчинения;
 - 4 — поселки городского типа, которые входят в состав горсовета;
 - 5 — поселки городского типа, которые входят в состав райсовета;
 - 6 — поселки городского типа, которые входят в состав райсовета в городе;
 - 7 — города, которые входят в состав горсовета;
 - 8 — сельсоветы, которые входят в состав райсовета;
 - 9 — сельсоветы, села, которые входят в состав райсовета города, горсовета.

Національний класифікатор України. Класифікація видів економічної діяльності (КВЕД)

"Класифікація видів економічної діяльності" (КВЕД) являється составляющей системы национальных классификаторов. Разрабатывание КВЭД осуществлено на базе международной статистической классификации видов деятельности Европейского Союза — Nomenclature of Activities European Community (NACE, Rev.1, mod.7).

КВЭД предназначено для использования органами государственного управления, финансовыми органами и органами статистики. Код вида деятельности является одним из основных показателей стратификации статистической совокупности для организации сплошных и выборочных статистических наблюдений и основным его назначением является обеспечение:

- статистического учета предприятий и организации за видами экономической деятельности в Единственном государственном реестре предприятий и организации Украины (ЕГРПОУ);
- проведение статистических обследований экономической деятельности и анализа статистической информации на макроуровне (составление межотраслевого баланса производства и распределения товаров и услуги в соответствии с системой национальных счетов);
- сопоставление национальной статистической информации с международной путем применения единственной статистической терминологии, статистической единицы и принципов определения и изменения видов экономической деятельности предприятий и организаций.

Объектами классификации в КВЭД являются виды экономической деятельности статистической единицы (юридических лиц, отделенных подразделов юридических лиц, физических лиц — предпринимателей и тому подобное), которая на высших уровнях классификации группируется в отрасли.

Экономическая деятельность - это процесс сочетания действий, которые приводят к получению соответствующего набора продукции или услуги. Вид деятельности характеризуется использованием ресурсов, производственным процессом, выпуском продукции и предоставлением услуги. Классификация видов экономической деятельности построена так, что выделенные объекты охватывают все

или почти все виды деятельности, которые осуществляются в любой отрасли хозяйства.

КВЭД гармонизована с Классификацией видов экономической деятельности Статистической Комиссии Европейского Союза (НАСЕ) (Rev.1. модификация 7) на уровне классов, что дает возможность использовать ее для сравнения национальных статистических данных с данными Статистической Комиссии Европейского Союза без переходных ключей. Для обеспечения возможности сравнения национальных статистических данных с данными Международной стандартной отраслевой классификации видов экономической деятельности (ISIC) Организации Объединенной Нации в структуре КВЭД предоставлена графа "Код ISIC".

По структуре Классификация состоит из двух блоков: идентификация и названия. Блок идентификации имеет иерархическую систему классификации и систему кодирования с применением литерно-цифрового кода. Для выделения отрасли экономики (сельское хозяйство, металлургия, транспорт, строительство и тому подобное), которой принадлежат объекты классификации, введенные "Секция" и "Подсекция" обозначены прописными буквами латинского алфавита. Более глубокая детализация группировки КВЭД (раздел, группа, класс, подкласс) обозначается цифровыми кодами.

Общее кодовое обозначение объектов КВЭД имеет такой вид:

Y YY XX.XX.X

где	Y	- секция (от А к Q)
	YY	- подсекция (используется только в секции "С" (от "СА" к "СВ") и "D" (от "DA" к "DN"))
	XX	- раздел
	XX.X	- группа
	XX.XX	- класс
	XX.XX.X	- подкласс

Уровень подкласса имеет значение от "0" к "9". Значение "0" предоставляется подклассу в случае отсутствия расширения международного уровня класса на национальном уровне. Значения от "1" к "8" предоставляются подклассам в случае детализации

международного уровня класса. Значение "9" всегда означает другую группировку.

Пример: кодовое обозначение вида экономической деятельности D DA 15.11.0 имеет такое содержание:

D	- секция "Обробна промисловість" - подсекция "Харчова промисловість та перероблення сільськогосподарських продуктів"
DA	перероблення сільськогосподарських продуктів"
15	- раздел "Харчова промисловість"
15.1	- группа "М'ясна промисловість"
15.11	- класс "Виробництво м'яса та субпродуктів"
15.11.0	- подкласс "Виробництво м'яса та субпродуктів"

Класифікатор топографічної інформації, яка відображається на топографічних планах масштабі 1:5000, 1:2000, 1:1000, 1:500

"Класифікатор топографічної інформації, яка відображається на топографічних планах масштабів 1:5000, 1:2000, 1:1000, 1:500" єсть ведомственный классификатор. Он утверждён приказом Главного управления геодезии, картографии и кадастра, при Кабинете Министров Украины №25 от 9.03.2000 г.[35]

Классификатор топографической информации предназначен для использования в автоматизированной системе обработки топографической информации и служит для формализованного представления данных об элементах и объектах местности, которые отображаются на топографических планах масштабам 1:500 - 1:5 000.

По содержанию Классификатор являет собой систематизированную сводку кодовых обозначений элементов и объектов местности, а также признаков, которые характеризуют эти объекты при отображении сведений о местности на топографических планах.

Топографическая информация, которая включена в Классификатор, разделена на две связанные между собой части:

- информация непосредственно об элементах, объектах местности, которая вмещает сведения об основных признаках и постоянных свойствах, однозначно определяет объект в общей системе классификации;

- информация о переменных свойствах, которые характеризуют объект и его отношение к другим объектам.

В основу построения Классификатора положен иерархический метод классификации информации первого типа и фасетный метод классификации информации второго типа.

Вся информация первого типа на высшей ступени иерархии разделена на восемь классов за элементами содержания топографических планов. Развитие каждого класса в глубину и ширину неодинаковая и зависит от степени взаимосвязи объектов и выбранных признаков классификации.

На низшей степени классификации каждого элемента содержания находится совокупность однотипных элементарных объектов плана, которые являются минимальной величиной (единицей) данных о местности, которые трансформируются в процессе обработки картографической информации.

Каждый элементарный объект плана вмещает предварительно определенный набор характерных признаков (количественных и качественных), которые не использовались в качестве признака классификации. Количество этих признаков и их смысловое значение разные и зависят от типа объекта.

В целом система классификации элементов и объектов местности, которые отображаются на топографических планах, характеризуется такими показателями:

- максимальное количество уровней классификации (глубина) - 8;
- максимальное число классификационной группировки на одном уровне - 9.

Табл. 3.2.5 – Перечень наименований объектов классификации, классификационной группировки и их кодовых обозначений (фрагмент)

Код об'єкта	Найменування класифікаційного угруповання	Найменування об'єкта класифікації	Коди ознак, які характеризують об'єкт класифікації
10000000	ГЕОДЕЗИЧНІ ПУНКТИ, ПОЗНАЧКИ ВИСОТ		
11000000	Геодезичні пункти		
11100000		Астрономічні пункти	09
11200000		Пункти державної геодезичної мережі	04, 09, 20, 21, 37, 68, 69
11300000	Пункти геодезичних мереж згущення та точки знімальної мережі		
11310000		Пункти геодезичних мереж згущення	04, 09, 20, 21, 37, 68, 69
11320000		Точки планових знімальних мереж тривалого закріплення на місцевості	04, 20, 37, 69, 119
11330000		Точки планових знімальних мереж тимчасового закріплення на місцевості	04, 20, 37, 69, 119
11340000		Точки планових знімальних мереж на рогах капітальних будівель	119

Табл. 3.2.6 – Перечень признаков, которые характеризуют объекты классификации, смысловых значений признаков и их кодовых обозначений (фрагмент)

Код ознаки	Найменування ознаки, яка характеризує об'єкт класифікації	Смислове значення ознаки характеристик	Коди смислового значення	Примітка	Код ознаки
55	Матеріал покриття	асфальт (асфальто-бетон)	1	6	55
		бетон	11		
		бітумо-мінеральна суміш	10		
		цементобетон	2		
		булижник	3		
		брущатка	4		
		гравій	5		
		камінь колотий	6		
		клинкер	7		
		шлак	8		
		щебінь	9		

Классификатор базовых наборов геопространственных данных

За основу создания базовых наборов геопространственных данных целесообразнее использовать первичную модель данных, которые создаются при цифровом картографировании. Эта целесообразность обуславливается возможностью сохранения точности представления формы, размеров и позиционирования пространственных объектов, которое определяется технологией цифрового картографирования.

Существующая технология цифрового картографирования ориентирована: а) на издание бумажной топографической карты, б) в определенном масштабе, в) в условных обозначениях. В этой связи возникает расхождение в представлении пространственных данных

цифрового картографирования с моделью данных в геоинформационной системе.

Поэтому данные цифрового картографирования нуждаются в инструментах трансформации их в информационную среду геоинформационной системы. С этой целью базовые наборы геопространственных данных следуют создавать путем корректной трансформации данных цифрового картографирования без потери точности определения объектов на основании Классификатора базовых наборов геопространственных данных.

Классификатор базовых наборов геопространственных данных [42] предназначен для использования в геоинформационной системе. В нем представлены 46 классов пространственных объектов, которые в свою очередь содержат 248 подтипов, которые являются основой для создания базовых слоев геоинформационной системы.

Таблица 3.2.7 - Фрагмент классификатора базовых наборов геопространственных данных

Код класу, підтипу	Клас просторових об'єктів	Підтип класу просторових об'єктів	Тип геометрії
61 300 000	Road_L Лінійні елементи проїжджих частин вулиць, тротуарів, покриття		Лінійний
61 341 100		Бордюр проїжджої частини	
61 342 100		Край проїжджої частини без бордюру	
61 346 000		Межі зміни покриття	
61 347 000		Вісь вулиці	
61 343 100		Край тротуару з покриттям без бордюру	
61 343 200		Бордюр тротуару з покриттям	
61 344 100		Край тротуару (доріжки) без покриття	

Классификатор базовых наборов геопространственных данных является та ким, который можно развивать.

3.2.4 Контрольные вопросы и задания для самостоятельной работы

- 1) Дайте определение классификации. Какую цель преследует классификация?
- 2) Какие сложились подходы к классификации?
- 3) Какое значение имеют классификации для геоинформационных систем?
- 4) Опишите основные термины и определения классификации.
- 5) Опишите основные методы классификации.
- 6) Опишите основные методы кодирования объектов классификации.
- 7) Приведите общую характеристику Единой системы классификации и кодирования технико-экономической информации.
- 8) Дайте характеристику пяти подходов при решении классификационных задач в геоинформационных системах.
- 9) Постройте свою классификацию объектов реального мира, которые вы видите.

Раздел 3.3

СИСТЕМЫ КООРДИНАТ И КАРТОГРАФИЧЕСКИЕ ПРОЕКЦИИ

3.3.1 Значение координат в геоинформационных системах

Рассмотрение компьютерных моделей географических объектов позволяет выделить сущность представления реального мира в ГИС: модели географических объектов и их пространственные свойства – местоположение, форма, размеры, пространственные отношения – представляются посредством координат. Следовательно, и обработка географических данных в свою очередь связана с обработкой координат географических объектов. Таким образом, координаты являются базовым компонентом в моделях географических объектов.

Для определения местоположения объектов на Земле используются системы координат, введенные для реального пространства Земли. Определение систем координат для реального пространства Земли имеет свою специфику и является предметом рассмотрения геодезической науки.

Область геоинформационной науки не только интегрирует решения, выработанные геодезической наукой и картографией в этом направлении, но и развивают их в системной информационной среде. Геоинформационные системы содержат решения, реализованные в программном продукте, связанные с установлением системы отсчета, трансформированием данных из одной системы координат в другую, созданием топологических моделей, анализом данных при условии представления этих данных в единой системе координат, и другие. Для решения многих задач в ГИС становится ключевым процесс, который получил название *геореференция* (Geo-referencing) – установление связей модели географических объектов с системой координат реального пространства Земли и картографической проекцией.

Поэтому для работы в геоинформационных системах требуются знания о базовых положениях систем координат в реальном земном пространстве и картографических проекциях.

3.3.2 Земные сферы и сфероиды

Физическая поверхность планеты Земля имеет сложную форму. *Физическую модель* фигуры Земли представляет *геоид* (Geoid) – уровенная поверхность, к которой отвесные линии всюду перпендикулярны и которая проходит через точку начала отсчета высот на среднем уровне океана. Высоты геоида относительно земного эллипсоида изменяются от +85 м. до -107 м. Ученые из высшей технической школы Цюриха и геодезического института Мюнхена в середине 80-х годов построили модель физической поверхности Земли в виде деформированной объемной фигуры (Рис.3.3.1) на основании анализа спутниковых данных о силе земного притяжения в различных точках планеты [Известия, 14.02.1987]. Спутниковая технология показала несколько эллиптических отклонений, например, Южный полюс ближе к экватору, чем Северный полюс.

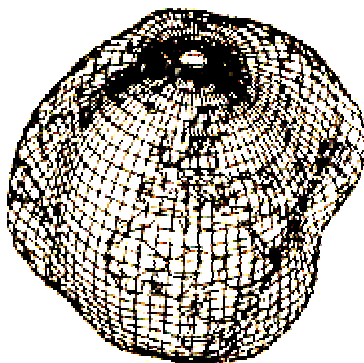


Рис. 3.3.1 – Геоид

Сложную поверхность планеты Земля аппроксимируют двумя *геометрическими моделями*: 1) сферой (Sphere), 2) сфероидом (Spheroid) - эллипсоидом (Ellipsoid) вращения, сжатым с полюсов. Выбор геометрической модели Земли зависит от цели решаемой задачи и необходимой точности результата вычислений.

Представление Земли в виде *сферы* упрощает математические вычисления. Сфера основана на круге и описывается одним параметром – радиусом Земли. Земная сфера применяется в тех случаях, когда различие между сферой и сфероидом не существенно для точности результатов решения задачи. Например, при

мелкомасштабном картографировании масштаба 1:5000000 и меньше, а также в ряде других случаев практики земной сфероид заменяют земной сферой (Earth's Sphere, Terrestrial Globe).

Представление Земли в виде *сфероида* связано со сложными математическими вычислениями и служит для вычислений геодезических широт, долгот, азимутов, длин, площадей, расчетов картографических проекций и решения других практических задач. Фигура Земли аппроксимируется сфероидом для отображения поверхности Земли на картах масштаба 1:1000000 и более крупных. Сфероид основан на эллипсе. Форма эллипса определена двумя радиусами - большой полуосью a и малой полуосью b .

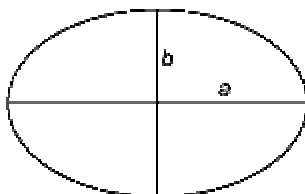


Рис. 3.3.2 - Полуоси сфероида.

Сфероид определяют два параметра: значения большой полуоси a и малой полуоси b или значение большой полуоси a и сжатие f .

Сжатие f вычисляется по формуле:

$$f = \frac{a - b}{a} \quad (3.3.1)$$

Сжатие характеризует различие в длинах полуосей. Сжатие может изменяться от нуля до единицы. Значение сжатия равное нулю означает, что эти две оси равны и образуют сферу. Сжатие сфероида Земли - маленькая величина, выражаемая десятичным числом (приблизительно равна 0,003353). Сжатие обычно представляется аликвотной дробью $1 / (1 : f)$.

Другая характеристика, которая описывает форму сфероида подобно сжатию, есть квадрат эксцентриситета e^2 меридианного эллипса, который представлен формулой:

$$e^2 = \frac{a^2 - b^2}{a^2} \quad (3.3.2)$$

На основании геодезических измерений в разное время определен ряд сфероидов, которые представляют Землю. В таблице представлены некоторые сфероиды.

Табл. 3.3.1 – Параметры сфероидов

Название	a	b	f	Применение
Airy 1830	6377563.396		1:299.3249646	
Bessel 1841	6377397.155	6356078	1:299.1528128	Central Europe
Clarke 1866	6378206.4	6356583	1:294.9786982	GB
Clarke 1880	6378249.145		1:293.465	N. America, France
Everest 1830	6377276.345	6356075	1:300.8017	India, Burma
Fischer 1960 (Mercury)	6378166		1:298.3	
Fischer 1968	6378150		1:298.3	
Helmert 1907	6378200	6356818		Egypt
Heyford 1909	6378388	6356912	1:297.0	Finland
GRS 1967	6378160		1:298.247167427	
GRS 1975	6378140		1:298,257	
GRS 1980	6378137		1:298.257222101	Ukraine
Hough 1956	6378270		1:297.0	
International	6378388	6356772	1:297.0	
Krassovsky 1940	6378245	6356863	1:298.30003	USSR, E. Europe
South American 1969	6378160		1:298.25	South American
WGS60	6378165		1:298.3	
WGS66	6378145		1:298.25	
WGS72	6378135		1:298.26	
WGS 84 Ellipsoid	6378137,0	6356752	1:298.257223563	World Wide
ПЗ 90 параметры Земли 1990	6378136		1:298,25784	Russia

Общеземной эллипсоид (World Ellipsoid) аппроксимирует Землю в целом. Референц-эллипсоид (Reference Ellipsoid) – это локальный эллипсоид, принятый для обработки измерений и установления системы геодезических координат для определенного региона или одной страны, который лучше всего приспособлен к поверхности

этого региона или страны. Например, до недавнего времени, данные США, Латинской и Центральной Америки использовали сфероид, определенный Кларком (Clarke) в 1866. Большая полуось сфероид Кларка 1866 имеет длину 6378206,4 метра, а малая полуось – 6356583,8 метра. Новый стандартный сфероид для Северной Америки - Геодезическая опорная система GRS 1980 (Geodetic Reference System - 1980), радиусы которого составляют 6378137,0 и 6356752,31414 метра. В любом случае нужно знать, что замена сфероид, принятого для системы координат, изменяет все предварительно измеренные значения.

3.3.3 Геодезические даты

Геодезические даты, датумы (Datums, Geodetic datums) – это система отсчета, которая описывает форму и размеры Земли, начало, ориентацию и масштаб координатных систем, используемых для определения местоположения относительно Земли посредством координат. Даты фиксируют положение сфероид в теле Земли. Они определяют начало и ориентацию линий долгот и широт. Даты обеспечивают основу пространственной привязки для измерения местоположения на поверхности Земли.

Всякий раз, когда изменяются даты, значения координат данных изменяются. Например, один и тот же контрольный пункт в городе Редланд (Redland) Калифорнии имеет координаты в Североамериканских датах 1927 (NAD27):

широта $34^{\circ} 01' 43,72995''$, долгота $-117^{\circ} 12' 54,61539''$;

в Североамериканских датах 1983 (NAD83):

широта $34^{\circ} 01' 43,77884''$, долгота $-117^{\circ} 12' 57,75961''$;

Значения долготы отличаются примерно на 3 секунды, в то время как значения широты отличаются примерно до 0,05 секунд.

По охвату земного пространства даты подразделяются на геоцентрические (общеземные) и топоцентрические (местные).

Геоцентрические даты определяются общим земным сфероидом. В последние годы спутниковые данные обеспечили геодезистов новыми измерениями, чтобы определить лучше аппроксимирующий Землю сфероид. Развитыми и широко используемыми датами является Мировая геодезическая система 1984 (World Geodetic System of 1984 - WGS 1984). Она служит как основа для измерения местоположения во всем мире.

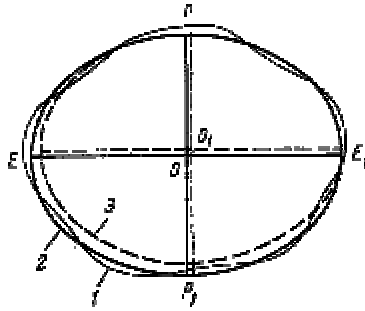


Рис. 3.3.3 – Референсные поверхности: 1) геоид, 2) сфероид общеземных дат, 3) сфероид локальных дат [43].

Топоцентрические даты определяют сфероид так, чтобы посадить его как можно ближе к поверхности Земли в специфической области. Начало системы координат местных дат не в центре Земли. Центр сфероидов местных дат смещен от центра Земли. Например, местные даты NAD 1927 разработаны так, чтобы посадить Северную Америку хорошо, в то время как Европейские даты (European Datum) ED 1950 были созданы для использования в Европе. Поскольку местные даты подгоняют их сфероид близко к специфической области на поверхности Земли, они не подходят для использования вне области, для которой они были разработаны.

По размерности координатных систем различают следующие типы дат:

- *Горизонтальные даты* определяют отношение между физической Землей и горизонтальными координатами, такими как широта и долгота;
- *Вертикальные даты* определяют уровенные поверхности. Одни вертикальные даты опираются на измерения уровня океана, другие – на гравиметрические измерения;
- *Комплексные даты* определяют вертикальные и горизонтальные системы. Например, Мировая Геодезическая Система (WGS-84) описывает также параметры вращения Земли и различные физические константы, такие как угловая скорость Земли и гравитационные постоянные Земли.

3.3.4 Системы координат

Для обработки данных в геоинформационных системах используются следующие типы систем координат:

- *системы координат связанные с моделью Земли* - системы пространственных прямоугольных координат, системы криволинейных координат, проецированные системы координат;
- *системы координат не связанные с моделью Земли* – неизвестные (Unknown) системы локальных координат.

Референционные системы координат – это такие системы координат, которые связаны с объектом в определенной системе отсчета.

3.3.4.1 Мировая геодезическая система координат WGS-84

Системой пространственных прямоугольных координат, связанной с моделью Земли, является Мировая геодезическая система WGS 84 (World Geodetic System 1984) [44]. Она имеет широкое применение в связи с использованием спутниковых данных, получаемых в глобальной системе позиционирования GPS.

Мировая геодезическая система WGS 84 была определена по спутниковым доплеровским измерениям в январе 1987 г. В последующие 1994 и 1996 годы WGS 84 была усовершенствована за счет дополнительных точных спутниковых, геодезических и гравиметрических измерений. Третья редакция Мировой геодезической системы WGS 84 имеет расхождение несколько сантиметров в координатах с Международной общеземной системой, которая поддерживается Международной службой вращения Земли IERS (International Earth Rotation Service).

Мировая геодезическая система координат WGS-84 - это общепринятая земная референционная система CTRS (Conventional Terrestrial Reference System), которая соответствует критериям IERS. Система координат WGS-84 есть геоцентрическая система. Центр масс определен для всей Земли, включая океаны и атмосферу, и соответствует геометрическому центру Эллипсоида WGS-84 с фундаментальными геодезическими постоянными, представленными в табл. :

Таблица 3.3.2 - Фундаментальные геодезические постоянные эллипсоида WGS-84

Параметр	Обозначение	Значение
Экваториальная полуось	a	6378137.0 m
Сжатие	f	1:298.257223563
Угловая скорость вращения Земли	ω	$7292115.0 \times 10^{-11}$ rad/s
Произведение гравитационной постоянной на массу Земли	GM	$3986004.418 \times 10^8 \text{ m}^3/\text{s}^2$
Скорость электромагнитных волн в вакууме	c	299792458 m/s

Ориентация Эллипсоида WGS-84 была изначально задана Международным бюро ориентации.

Координатная система WGS-84 является правосторонней ортогональной декартовой системой координат.

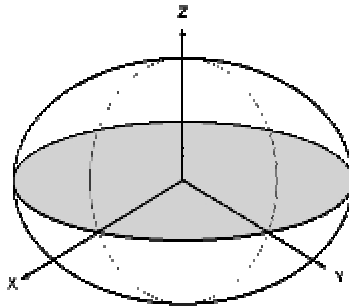


Рис. 3.3.4 - Мировая геодезическая система координат WGS-84

Формулы связи прямоугольных пространственных координат (x, y, z) с геодезическими координатами (φ, λ, h) :

$$x = (N + h) \cos \varphi \cos \lambda \quad (3.3.3)$$

$$y = (N + h) \cos \varphi \sin \lambda \quad (3.3.4)$$

$$z = ((b^2 / a^2)N + h) \sin \varphi \quad (3.3.5)$$

где N - радиус кривой первого вертикала вычисляется по формуле:

$$N = \frac{a}{(1 - e^2 \sin^2 \varphi)^{1/2}} \quad (3.3.6)$$

Формулы связи геодезических координат (φ, λ, h) с прямоугольными пространственными координатами (x, y, z) :

$$\sin \lambda = \frac{y}{\sqrt{x^2 + y^2}} \quad (3.3.7)$$

$$\sin \varphi = \frac{z}{\sqrt{x^2 + y^2 + z^2}} \quad (3.3.8)$$

$$\rho = \sqrt{x^2 + y^2 + z^2} \quad (3.3.9)$$

$$h = \rho - a \quad (3.3.10)$$

3.3.4.2 Системы криволинейных координат

Криволинейными координатами на референционной поверхности являются угловые величины "*широта*" (latitude) и "*долгота*" (longitude), которые определяются направлением нормали к референционной поверхности.

В зависимости от вида референционной поверхности различают следующие виды географических координат:

- *Географические координаты*, базирующиеся на математически определенной нормали к поверхности земной сферы.
- *Астрономические координаты*, базирующиеся на физически определенной отвесной линии к уровенной поверхности геоида.
- *Геодезические координаты*, базирующиеся на математически определенной нормали к поверхности земного сфероида.

Основанные на этом принципе системы координат соответственно называются:

- Географическая система координат,
- Астрономическая система координат,
- Геодезическая система координат.

Термины "географические координаты", "географическая долгота", "географическая долгота" имеют историческое происхождение. Они появились тогда, когда Землю принимали за шар, а координаты определяли астрономическими методами. Позже с определением физической и геометрических моделей фигуры Земли появились виды криволинейных координат, в которых термины "широта" и "долгота" сохранились, а термин "географические

координаты" иногда применяется в обобщенном представлении криволинейных координат.

Географическая система координат (Geographic Coordinate System - GCS) использует трехмерную сферическую поверхность, чтобы определить местоположения на Земле. GCS включает угловую единицу измерения, главный меридиан, экватор и даты, основанные на сфере.

Геодезическая система координат использует трехмерную сфероидическую поверхность, чтобы определить местоположения точек на Земле геодезическими координатами, которые в ГИС обозначаются φ , λ , h (в странах бывшего СССР геодезические координаты обозначаются В, L, Н соответственно).

Геодезическая широта φ точки – это угол между плоскостью экватора и нормалью к поверхности сфероида в точке.

Геодезическая долгота λ точки – это угол между плоскостью начального меридиана и плоскостью геодезического меридиана точки.

Геодезическая высота h точки – это длина отрезка нормали от сфероида до точки.

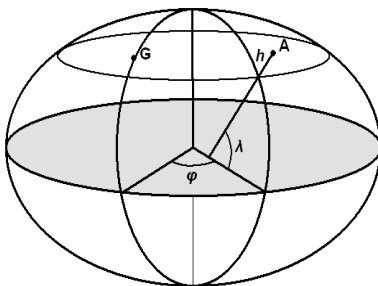


Рис. 3.3.4 – Геодезическая система координат

Кривыми координатными линиями являются меридианы и параллели. Меридиан есть координатная линия, во всех точках которой долгота имеет постоянное значение. Параллель есть координатная линия, для которой широта имеет постоянное значение. Эти координатные линии охватывают земной шар и формируют сеть параллелей и меридианов.

Линия нулевой широты или максимальная параллель называется экватором. Линия нулевой долготы называется главным меридианом. Для большинства географических систем координат главный меридиан проходит через Гринвич (Greenwich) в Англии.



Рис. 3.3.5 – Обозначение главного меридиана в Гринвичской обсерватории [45].

Начало сетки параллелей и меридианов (0,0) определено там, где пересекаются экватор и главный меридиан. Значения широты и долготы измеряются или в градусах, минутах, и секундах (degrees, minutes, and seconds - DMS), или в десятичных градусах. Значения широты изменяются относительно экватора в диапазоне от -90° в Южном полюсе до $+90^\circ$ в Северном полюсе. Значения долготы изменяются относительно главного меридиана от -180° при перемещении на запад до $+180^\circ$ при перемещении на восток. Долгота и широта могут иметь точные положения на поверхности земного шара.

Поскольку меридианы сходятся к полюсам, расстояние, представленное одним градусом долготы, не есть постоянная величина; оно уменьшается к полюсам. Так, например, на сфероиде Кларка (Clarke) 1866 один градус долготы в экваторе равняется 111.321 км. ($1^\circ - 111$ км., $1' - 1.85$ км., $1'' - 31$ м., $0.001'' - 3$ см.). На широте 60° один градус долготы равняется 55.802 км. Так как градусы широты и долготы не имеют стандартную длину, для точного

определения расстояния или площади требуются сложные математические зависимости.

3.3.4.3 Проецированные системы координат

Проецированными системами координат (Projected coordinate systems) называются любые системы координат, разработанные путем проецирования выпуклой поверхности Земли на плоскую двумерную поверхность типа бумажной карты или компьютерного экрана. Это двумерная система координат, которую используют для аппроксимации формы Земли на плоской поверхности.

Проецированная система координат представляет собой плоскую декартову систему прямоугольных координат (x , y), простую в применении. В отличие от географической системы координат, проецируемая система координат имеет постоянные длины, углы, и площади на двумерной плоскости. Проецированная система координат всегда основана на криволинейной системе координат, которая опирается на сферу или сфероид. Координаты проецированной системы координат вычисляются по координатам географической / геодезической системы координат.

3.3.5 Картографические проекции

3.3.5.1 Что представляет собой картографическая проекция?

Картографической проекцией (Map Projection) называется математическое преобразование трехмерной поверхности Земли, рассматриваемой как сфера или сфероид, на плоскость.

Многие картографические проекции используют идею проецирования, которую поясняет рис. 3.3.6. Поверхность Земли проецируется на поверхность проецирования. В последующем поверхность проецирования разворачивается в плоскость, для которой вводится проецированная система координат. В таких случаях поверхности проецирования называют разворачивающимися поверхностями.

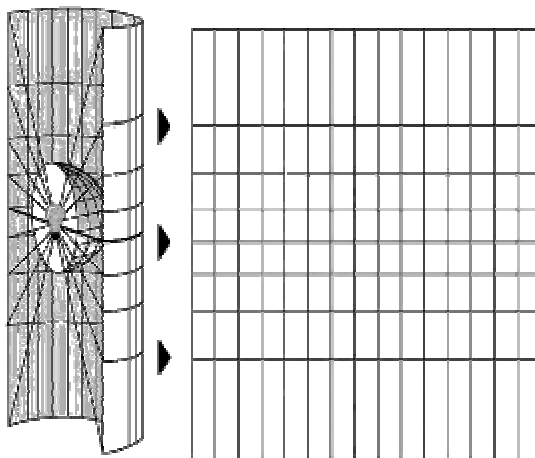


Рис. 3.3.6 – Иллюстрация проецирования поверхности Земли [43]

Картографическая проекция использует математические формулы, чтобы связать сфероидические/сферические координаты с плоскими плановыми координатами. В общем случае математическое преобразование представляется зависимостями:

$$x = f_1(\varphi, \lambda) \quad (3.3.11)$$

$$y = f_2(\varphi, \lambda) \quad (3.3.12)$$

В результате проецирования на плоскости получается искаженной сетка поверхности Земли. Это иллюстрация того, что сфера или сфероид не могут быть развернуты в любую плоскость без искажений. Представление поверхности Земли в двух измерениях приводит к искажениям формы, площади, расстояния или направления. Чтобы уменьшить искажения, трехмерные пространственные объекты должны плотно прилегать к поверхности проецирования.

Различные проекции вызывают различные типы искажений. Некоторые проекции разработаны, чтобы минимизировать искажения одной или двух характеристик данных. Проекция может представить площадь пространственного объекта, но изменить его форму.

Картографические проекции разработаны для определенных целей. Одна картографическая проекция может использоваться для крупномасштабных данных в ограниченной области, другая используется для мелкомасштабной карты мира. Картографические

проекции, разработанные для мелкомасштабных данных, обычно основаны на сферических, а не на сфероидальных системах координат.

3.3.5.2 Классификация проекций по типу искажений

По типу искажений картографические проекции делятся на конформные, равноплощадные, равнопромежуточные и азимутальные.

Конформные проекции

Конформные проекции (Conformal projections) сохраняют локальную форму. Чтобы сохранять индивидуальные углы, описывающие пространственные отношения, конформная проекция должна показать перпендикулярные линии географической сетки, пересекающиеся на карте под углом 90 градусов. Картографическая проекция сохранит все углы. Отрицательная сторона проекции заключается в том, что площадь, ограниченная серией дуг, может быть значительно искажена в процессе. Никакая картографическая проекция не может сохранять формы больших областей.

Равноплощадные проекции

Равноплощадные проекции (Equal area projections) сохраняют площадь отображенных пространственных объектов. Чтобы обеспечить это, другие свойства - форма, угол, и масштаб - искажаются. В равноплощадных проекциях меридианы и параллели не могут пересекаться под прямыми углами. В некоторых случаях, особенно на картах малых регионов, формы очевидно не искажены, и трудно отличить равноплощадную проекцию от конформной проекции без документирования или измерений.

Равнопромежуточные проекции

Равнопромежуточные проекции (Equidistant projections) сохраняют расстояния между определенными точками. Любой проекцией масштаб не поддерживается корректно везде в полной карте. Однако, в большинстве случаев имеются одна или большее число линий на карте, по которым масштаб поддерживается правильно. Большинство проекций, сохраняющих расстояния, имеют одну или большее количество линий, для которых длина линии на карте является той же самой длиной (в масштабе карты), как такой же самой линии на

земном шаре. Такие расстояния считаются истинными и отвечают истинному масштабу между двумя точками на карте.

Азимутальные проекции

Самый короткий маршрут между двумя точками на кривой поверхности Земли проходит по сферическому эквиваленту прямой линии на плоской поверхности. Это - большой круг, проходящий через две точки. Проекции истинного направления (True-direction projections) или азимутальные проекции содержат некоторые дуги большого круга, представляющие направления или азимуты всех точек на карте правильно относительно центра. Некоторые проекции истинного направления могут быть также конформными, равноплощадными или равнопромежуточными.

3.3.5.3 Классификация проекций по типу поверхности проецирования

Картографические проекции классифицированы в соответствии с использованной поверхностью проецирования - *конической, цилиндрической* или *плоской* (Рис. 3.3.7). Она может быть выровнена без растягивания их поверхности. Картографическая проекция проецирует местоположения на поверхности сфероида к соответствующим позициям на плоской поверхности, используя математические зависимости.

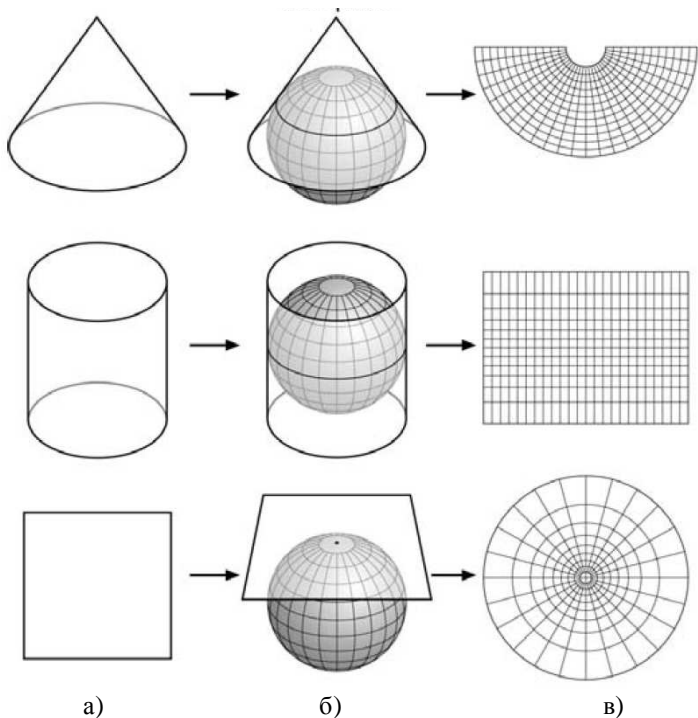


Рис. 3.3.7 - Коническая, цилиндрическая, плоская поверхности проецирования: а) исходные формы, б) развертывающиеся поверхности, в) проекции [46]

Поверхности проецирования - конусы, цилиндры и плоскости могут принимать *нормальное* (Normal), *поперечное* (Transverse) или *наклонное* (Oblique) положение относительно оси вращения сферы или сфероида.

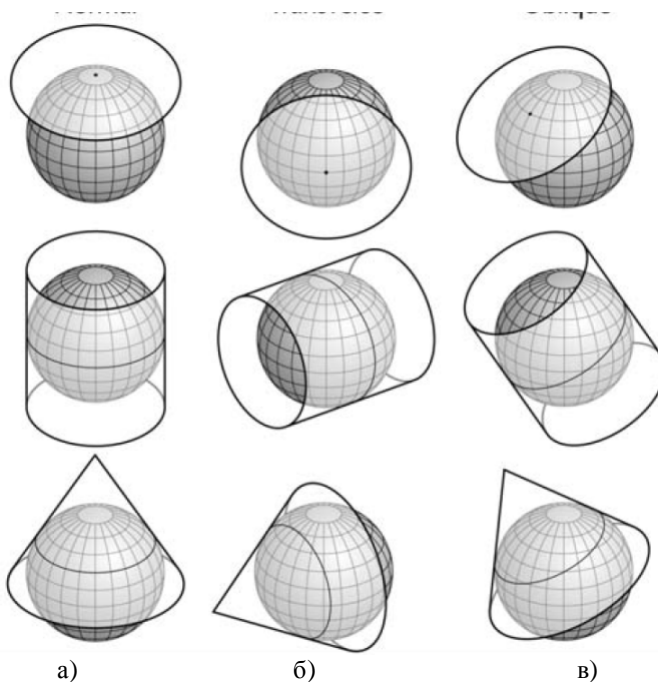


Рис. 3.3.8 - Положение поверхности проецирования:
 а) нормальное, б) поперечное, в) наклонное [46]

Поверхности проецирования - конусы, цилиндры и плоскости могут быть *касательными*, (Tangential) к сфере, сфероиду или секущими (Secant) сферу, сфероид. Контакт двух поверхностей происходит в точке, линии касания или по линии сечения. Независимо от того, что контакт касательный или секущий, точки или линии контакта являются существенными, так как они определяют размещения нулевого искажения. Линии истинного масштаба называют *стандартными линиями*. Вообще, искажение увеличивается с увеличением расстояния от точки контакта.

Конические проекции

Конические проекции (Conic projections) используют конус как поверхность проецирования. Наиболее простая коническая проекция -

касательная к земному шару по параллели. Эта линия является стандартной параллелью. Меридианы проецируются на коническую поверхность, соединяясь в вершине или точке конуса. Параллели проецируются на конус как кольца. Конус затем разрезается по любому меридиану и разворачивается в плоскость. Заключительная коническая проекция имеет прямо сходящиеся линии для меридианов и концентрические дуги окружностей для параллелей. Противоположный линии разреза меридиан становится центральным меридианом.

С удалением от стандартной параллели искажения становятся большими. Конические проекции используются для зон средних широт, которые имеют восточно-западную ориентацию, и не используются для полярных областей.

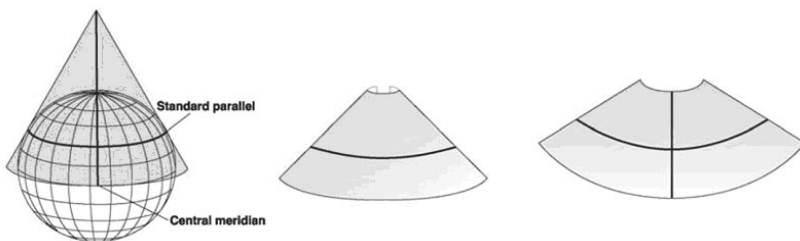


Рис. 3.3.9 - Касательная коническая проекция [43]

Секущие конические проекции определены двумя стандартными параллелями. Модель искажений для секущих проекций разная между стандартными параллелями и за ними. Секущая проекция имеет меньшие искажения, чем касательная проекция. Если ось конуса не совмещают с полярной осью земного шара, такие типы проекций называются перспективными.

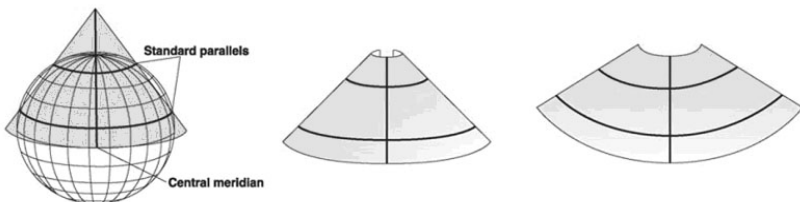


Рис. – 3.3.10- Секущая коническая проекция [43]

Представление географических пространственных объектов зависит от интервала между параллелями. Когда параллели равномерно расположены, такая проекция является

равнопромежуточной, но ни конформной, ни равноплощадной. Пример этого типа проекции - равнопромежуточная коническая проекция. Для малых областей полные искажения минимальны. Примером конформной конической проекции является проекция Ламберта (Lambert), а конической равноплощадной проекции проекция Альберта (Alber).

Цилиндрические проекции

Цилиндрические проекции (Cylindrical projections) используют цилиндр как поверхность проецирования. Подобно коническим проекциям, цилиндрические проекции могут также быть секущими или касательными. Ось цилиндра может располагаться по оси вращения эллипсоида, в плоскости экватора, наклонно к оси вращения эллипсоида.

При размещении оси цилиндра по оси вращения эллипсоида создается *нормальная* цилиндрическая проекция (Normal cylindrical projections). Примером является проекция Меркатора (Mercator), у которой экватор является ее линией касания. Сечения геометрически проецируются на цилиндрическую поверхность, а параллели математически проецируются. Этим создается координатная сетка с углами 90 градусов. Цилиндр разрезается по любому меридиану, чтобы построить конечную цилиндрическую проекцию. На плоскости меридианы равномерно расположены, в то время как интервал между параллельными линиями широты увеличивается к полюсам. Это проекция конформная, которая показывает истинное направление по прямым линиям. В проекции Меркатора локсодромы (линии постоянного направления) являются прямыми линиями, но большинство больших кругов не являются таковыми.

При размещении оси цилиндра в плоскости экватора эллипсоида создается *поперечная* цилиндрическая проекция (Transverse cylindrical projections). Поперечные цилиндрические проекции типа Трансверсальной Меркатора используют меридиан как тангенциальный контакт или линии параллельные меридианам как линии сечения. Стандартные линии в таком случае проходят от северного до южного полюса, вдоль которых масштаб является истинным.

При размещении оси цилиндра наклонно к оси вращения эллипсоида создается *наклонная* (косая) цилиндрическая проекция (Oblique cylindrical projections). Наклонные цилиндры вращаются вокруг линии большого круга, зафиксированной между экватором и

меридианами. В этих проекциях большинство меридианов и параллелей не являются прямыми линиями.

Во всех цилиндрических проекциях линия касания или линии сечения не имеют искажений и таким образом являются линиями равноудаленности. Другие географические свойства изменяется в соответствии со спецификой проекции.

Плоские проекции

Плоские проекции (Planar projections) используют плоскость как поверхность проецирования. Плоская проекция также известна как азимутальная проекция или зенитальная проекция. Плоскость может быть касательной к земному шару в одной точке, или может быть секущей. Точкой касания может быть Полярный Полос, Южный Полос, точка на экваторе или любая точка между ними. Возможные положения поверхности проецирования: полярное, экваториальное и наклонное (косое).

Полярное положение поверхности проецирования - самая простая форма. Параллели представляются концентрическими кругами с центром в полюсе, а меридианы - прямыми линиями, которые пересекаются с их истинными углами ориентации в полюсе. В других положениях, плоские проекции будут иметь углы координатной сетки 90 градусов в фокусе. Направления из фокуса точны. Большие круги, проходящие через фокус, представлены прямыми линиями; таким образом, кратчайшее расстояние от центра до любой другой точки на карте является прямой линией. Модели искажений площади и формы являются круговыми относительно фокуса. По этой причине азимутальные проекции приспособлены к круговым областям лучше, чем к прямоугольным областям. Плоские проекции используются наиболее часто для отображения полярных регионов.

Некоторые плоские проекции рассматривают поверхностные данные из определенной точки в пространстве. Точка зрения определяет, как сферические данные проецируются на плоскую поверхность. Центром перспективы может быть центр Земли, или поверхностная точка прямо противоположная фокусу, или удаленная точка вне земного шара.

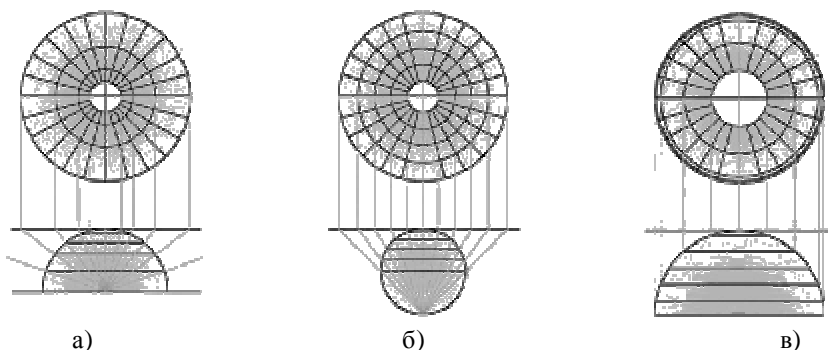


Рис. 3.3.11 - Азимутальные проекции: а) гномоническая, б) стереографическая) ортографическая [43]

По различным перспективам азимутальные проекции с полярным положением поверхности проецирования подразделяются на три вида:

- *Гномоническая* проекция рассматривает поверхностные данные из центра Земли;
- *Стереографическая* проекция рассматривает их от полюса до полюса;
- *Ортографическая* проекция рассматривает Землю из точки в бесконечности.

Различия в перспективе определяют количество искажений в направлении к экватору.

3.3.5.4. Другие проекции

Проекции, обсужденные выше концептуально, созданы проецированием одной геометрической формы (сферы, сфероида) на другую (конус, цилиндр или плоскость). Однако, многие другие проекции не связаны непосредственно с конусом, цилиндром или плоскостью.

Модифицированные проекции – это переработанные версии других проекций (например, Пространственная наклонная Меркатора есть модификация проекции Меркатора). Эти модификации разработаны, чтобы уменьшить искажения, часто включением стандартных дополнительных линий или изменением модели искажений.

Псевдопроекции имеют некоторые характеристики проекций другого класса. Например, синусоидальная проекция названа псевдоцилиндрической проекцией, потому что все параллели прямые и параллельны, и все меридианы равномерно расположены. Однако, это не истинно цилиндрическая проекция, потому что все меридианы кроме центрального меридиана изогнуты. Это приводит к карте Земли, имеющей овальную форму вместо прямоугольной формы.

3.3.5.5. Параметры картографических проекций

Картографическая проекция сама не определяет проецируемую систему координат. Можно заявить, что набор данных находится в проекции Трансверсальной Меркатора, но этой информации не достаточно. Где находится центр проекции? Использовался ли масштабный коэффициент? Без того, чтобы знать точные значения параметров проекции, набор данных не может проецироваться.

Каждая картографическая проекция имеет набор параметров, который необходимо задать. Параметры определяют начало координат и настраивают проекцию для области, представляющей интерес. Угловые параметры используют единицы географической системы координат, в то время как линейные параметры используют единицы проецируемой системы координат.

Линейные параметры

Восточное смещение координат (False easting) - линейное значение, применяемое к началу x правой системы координат.

Северное смещение координат (False northing) - линейное значение, применяемое к началу y правой системы координат.

Восточное и северное смещение координат применяются для того, чтобы все значения x или y были положительными. Параметры восточного и северного смещения координат можно также использовать, чтобы уменьшить диапазон x или y значений координат. Например, если все значения y большие, чем пять миллионов метров, можно применять северное смещение координат на 5000000 м.

Масштабный коэффициент - безразмерное значение, применяемое к центру или линии картографической проекции. Масштабный коэффициент - обычно несколько меньше чем единица. Система координат UTM, которая использует поперечную проекцию Меркатора, имеет масштабный коэффициент 0,9996 вдоль центрального меридиана проекции. Это создает две стандартные линии приблизительно на расстоянии 180 километров от центрального

меридиана, для которых масштабный коэффициент равняется 1,0, и в окрестности которых уменьшаются искажения проекции.

Угловые параметры

В одноточечных проекциях используются:

- *азимут* - измеряется от северного направления центрального меридиана;
- *долгота начала координат* определяет начало *x*-координат; параметры центральный меридиан и долгота начала координат синонимичны;
- *широта начала координат* определяет начало *y*-координат.

В двухточечной равнопромежуточной проекции Хотина (Hotine) (двухточечный центр и центр определения азимута) и косоугольной проекции Меркатора используются:

- *долгота центра*, чтобы определить начало *x*-координат;
- *широта центра*, чтобы определить начало *y*-координат.

В двухточечной равнопромежуточной проекции и в косоугольной проекции Хотина- Меркатора используются четыре параметра:

- *долгота первой точки*,
- *широта первой точки*,
- *долгота второй точки*,
- *широта второй точки*.

Они определяют две географические точки, которые фиксируют центральную ось проекции.

3.3.5.6 Примеры картографических проекций

Программные продукты ArcView GIS – ArcGIS включают инструменты для работы с 66 картографическими проекциями, которые описаны в руководстве [43]. Здесь приведена информация только о двух распространенных проекциях.

Проекция Гаусса-Крюгера (Gauss-Krueger)

Историческая справка

В 1569 году фламандский картограф Герард Меркатор (Mercator) разработал равноугольную цилиндрическую экваториальную проекцию, в которой площади искажались к полюсам. В 1772 году Йоганн Ламберт (Johann Heinrich Lambert) модифицировал

экваториальную проекцию Меркатора в трансверсальную форму с целью минимизации искажений. В 1822 году Йоганн Карл Фридрих Гаусс (Johann Carl Friedrich Gauss) выполнил анализ трансверсальной проекции, а в 1912 году Луис Крюгер (Louis Kruger) разработал формулы для эллипсоида. В Европе конформная проекция получила название Гаусса-Крюгера, а в США – трансверсальная Меркатора.

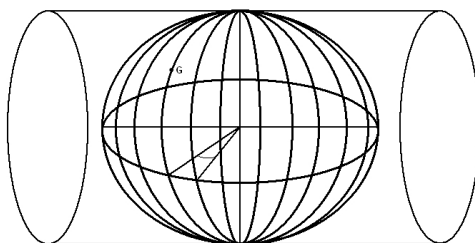


Рис. 3.3.12 - Проекция Гаусса-Крюгера

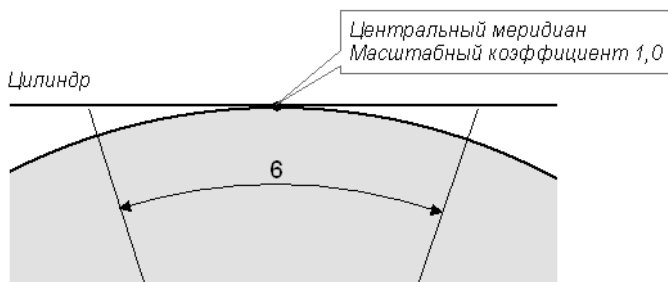


Рис. 3.3.13 – Касательный цилиндр проекции Гаусса-Крюгера

Описание

Проекция Гаусса-Крюгера является частным случаем проекции Transverse Mercator. Эта проекция подобна трансверсальной проекции Меркатора за исключением того, что цилиндр касательный меридиана вместо экватора. Проецирование выполняется по зоне, которая ограничена двумя меридианами. Центральный меридиан зоны расположен в области, которая будет проектируемой. Это центрирование минимизирует искажение всех свойств в этой области. Результат - конформная проекция, которая не хранит верные

направления.

Система координат Гаусса-Крюгера (ГК) основана на проекции Гаусса-Крюгера.

Метод проецирования

Цилиндрическая проекция с центральным меридианом, помещенным в специфическую область.

Линия контакта

Любой единственный меридиан для тангенциальной проекции.

Прямые линии сетки меридианов и параллелей

Экватор и центральный меридиан.

Свойства

Форма

Конформная проекция. Поддерживается малая форма. Форма большей области все более и более искажена с удалением от центрального меридиана.

Площадь

Искажение увеличивается с удалением от центрального меридиана.

Направление

Локальные углы точны всюду.

Расстояние

Точный масштаб вдоль центрального меридиана, если масштабный множитель равняется 1.0. Если масштабный множитель меньше чем 1.0, то есть две прямые линии, которые имеют точный масштаб, равноудаленные в каждую сторону от центрального меридиана.

Ограничение

Данные относительно сфероида в пределах 90 градусов от центрального меридиана не могут проецироваться. Фактически экстенд на сфероиде должен быть ограниченный от 10 до 12 градусов в обе стороны от центрального меридиана. В этом диапазоне обратное проецирование не возможно. При использовании сферы нет этих ограничений.

Использование и дополнения

Система координат Гаусса-Крюгера. Глобус разделяется на зоны шириной шесть градусов. Каждая зона имеет масштабный множитель 1.0 и смещение начала абсциссы 500000 метров. Центральный меридиан зоны 1 расположенный в 3° В. Некоторые регионы также добавляют зональный номер умноженный на один миллион к смещению

начала абсциссы на 500000. Так зона 5 может иметь смещение начала абсциссы в координате на 500000, или 5500000 метров. Эта проекция лучше подходит для континентальных массивов, которые вытянуты в направлении север-юг. Топографическое картографирование в бывшем Советском Союзе, Германии, Албании, Австрии.

Универсальная трансверсальная проекция Меркатора

В геоинформационных системах широко применяется Универсальная трансверсальная проекция Меркатора (Universal Transverse Mercator - UTM). С 1977 г. UTM используется Геологической съемкой США для точного картографирования.

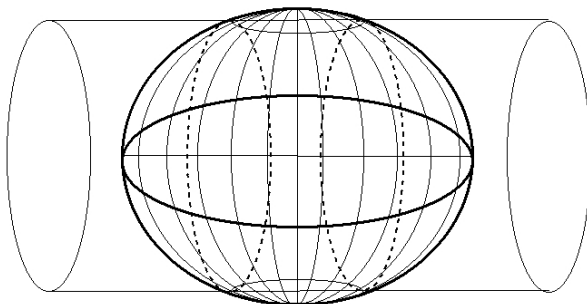


Рис. 3.3.14 - Проекция UTM

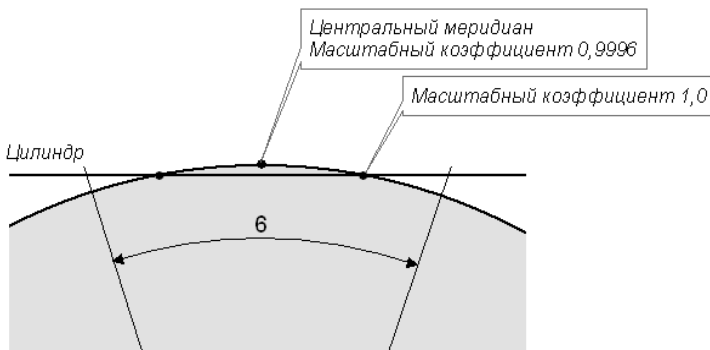


Рис. 3.3.15 - Секущий цилиндр проекции UTM

Описание

Универсальная система трансверсальной проекции Меркатора - специализированное приложение трансверсальной проекции Меркатора. Проекция UTM получается двойным проецированием - эллипсоида на шар, а затем шара на плоскость в проекции Меркатора. UTM - секущая проекция. Две параллельных линии равноудаленные от центрального меридиана.

Земной шар разделен на 60 северных и южных зон, каждая охватывает шесть градусов долготы. Каждая зона имеет собственный центральный меридиан. Центральный меридиан зоны 1 UTM - в 177° З. Зоны 1С и 1Ю начинаются от 180° З.

Пределы каждой зоны - 84° С и 80° Ю с подразделением на северную и южную зону, которые встречаются на экваторе. Полярная область использует универсальную полярную стереографическую систему координаты.

Начало для каждой зоны - ее центральный меридиан и экватор. Чтобы исключить негативную координату, система координат изменяет величину координаты в начале. Величина, присвоенная центральному меридиану, – смещение начала абсциссы, и величина, назначенная экватору, – смещение начала ординаты. Применяется смещение начала координаты на запад на 500000 метров. Северная зона имеет смещение начала ординаты 0, в то время как южная зона имеет смещение начала ординаты 10000000 метров.

Метод проецирования

Цилиндрическая проекция с центральным меридианом, размещенным в особенной области.

Линия контакта

Две линии параллельные центральному меридиану зоны UTM и удаленные от него приблизительно на 180 км с каждой стороны.

Прямая линия сетки меридианов и параллели

Центральный меридиан и экватор.

Свойства

Форма

Конформная проекция. Точное представление малой формы. Минимальное искажение больших форм в пределах зоны.

Площадь

Минимальное искажение в пределах каждой зоны UTM.

Направление

Локальные углы верные.

Расстояние

Масштаб постоянный по центральному меридиану, но в масштабном коэффициенте 0.9996, чтобы снизить боковое искажение в пределах каждой зоны. С этим масштабным коэффициентом линия, которая находится в 180 км к востоку и к западу и параллельны к центральному меридиану, имеет масштабный коэффициент 1,0.

Ограничение

Проецируемая для масштабной погрешности, меньшей 0,1 процента в пределах каждой зоны. Погрешность и искажение увеличиваются для области, которая перекрывает больше, чем одну зону UTM. UTM не разработана для области, которая перекрывает несколько зон.

Данные на сфероиде или эллипсоиде за 90 градусами от центрального меридиана не могут проектироваться. Фактически экстенд на сфероиде или эллипсоиде должен быть ограниченный 15-20 градусами с обеих сторон центрального меридиана. За этим диапазоном в поперечной проекции Меркатора проецируемые данных не имеют обратного проецирования. Данные относительно сферы не имеют этих ограничений.

Использование и дополнения

Используется в Соединенных Штатах для топографических листов масштаба 1:100000. Много стран используют локальные зоны UTM, основанные на применяемых официальных географических системах координат. Крупномасштабное топографическое картографирование в бывшем Советском Союзе.

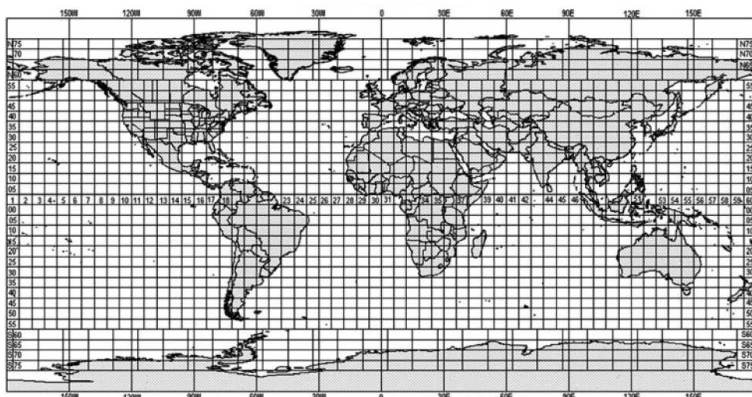


Рис. 3.3.16 – Покрытие земной поверхности зонами UTM [46].

3.3.7 Географические преобразования

Перемещение данных между системами координат включает преобразования между географическими системами координат. Поскольку географические системы координат содержат даты, которые основаны на сфероиде, географическое преобразование также изменяет подстилающий сфероид.

Известные методы имеют различные уровни точности и диапазоны для преобразования дат. Точность специфического преобразования может располагаться от сантиметров до метров в зависимости от метода, качества и числа контрольных точек, располагаемых, чтобы определить параметры преобразования.

Географическое преобразование конвертирует географические координаты (широты, долготы). Некоторые методы конвертируют географические координаты в геоцентрические (X , Y , Z) координаты, преобразовывают X , Y , Z координаты, и конвертируют новые значения обратно в географические координаты.

3.3.7.1 Метод с тремя параметрами

Самым простым методом преобразования дат является геоцентрическое преобразование или метод с тремя параметрами.

Геоцентрическое преобразование моделирует различия между двумя датами в системе координат X, Y, Z. Одни даты определены их центром в 0,0,0. Центр других дат определен на некотором расстоянии (Δx , Δy , Δz). Эти три параметра являются линейными сдвигами и всегда выражаются в метрах.

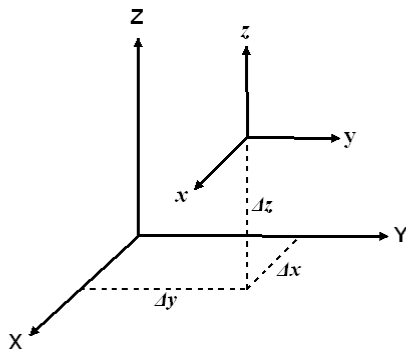


Рис. 3.3.17 - Метод с тремя параметрами

Обычно параметры преобразования определены как продвижение "от" местных дат "к" датам WGS 1984 или другим геоцентрическим датам. Геоцентрическое преобразование выполняется по формуле:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} + \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (3.3.13)$$

3.3.6.2 Метод с семью параметрами

Более комплексное и точное преобразование выполняется путем добавления трех линейных сдвигов (Δx , Δy , Δz) и еще четырех параметров к геоцентрическому преобразованию: три угловых поворота вокруг каждой оси (r_x, r_y, r_z) и масштабного коэффициента. Географическое преобразование выполняется по формуле:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} + (1 + s) * \begin{bmatrix} 1 & r_z & -r_y \\ -r_z & 1 & r_x \\ r_y & -r_x & 1 \end{bmatrix} * \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (3.3.14)$$

Значения поворота определяются различными способами - по часовой стрелке или против часовой стрелки.

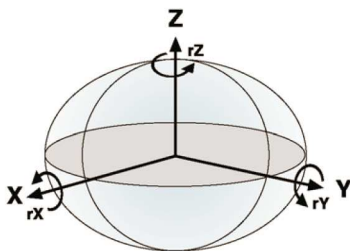


Рис. 3.3.17 - Направления угловых поворотов

3.3.6.3 Метод Молоденского

Метод Молоденского конвертирует данные непосредственно между двумя географическими системами координат фактически без преобразования в систему X, Y, Z. Метод Молоденского требует трех сдвигов (Δx , Δy , Δz) и различий между большими полуосями (Δa) и сжатия (Δf) этих двух сфероидов. Процессор проекции автоматически вычисляет различия сфероида согласно принятым датам, используя зависимости:

$$(M + h)\Delta\varphi = -\sin\varphi\cos\lambda - \sin\varphi\sin\lambda\Delta y + \cos\varphi\Delta z + \frac{e^2\sin\varphi\cos\varphi}{(1 - e^2\sin^2\varphi)^{1/2}}\Delta a \quad (3.3.15)$$

$$+ \sin\varphi\cos\varphi\left(M\frac{a}{b} + N\frac{b}{a}\right)\Delta f$$

$$(N + h)\cos\varphi\Delta\lambda = -\sin\lambda\Delta y \quad (3.3.16)$$

$$\begin{aligned} \Delta h = & \cos \varphi \cos \lambda \Delta x + \cos \varphi \sin \lambda \Delta y \\ & + \sin \varphi \Delta z - (1 - e^2 \sin^2 \varphi)^{1/2} \Delta a \\ & + \frac{a(1-f)}{(1 - e^2 \sin^2 \varphi)^{1/2}} \sin^2 \varphi \Delta f \end{aligned} \quad (3.3.17)$$

где

- h - высота эллипсоида;
- j - широта;
- l - долгота;
- a - большая полуось сфероида;
- b - малая полуось сфероида;
- f - сжатие сфероида;
- e - эксцентриситет сфероида;
- M, N - радиусы кривизны - меридианальный и первого вертикала соответственно в данной широте:

$$M = \frac{a(1 - e^2)}{(1 - e^2 \sin^2 \varphi)^{3/2}} \quad (3.3.18)$$

$$N = \frac{a}{(1 - e^2 \sin^2 \varphi)^{1/2}} \quad (3.3.19)$$

3.3.8 Контрольные вопросы и задания для самостоятельной работы

- 1) Каково значение координат в геоинформационных системах?
- 2) Приведите общие характеристики физических и геометрических моделей фигуры Земли.
- 3) Что представляют собой геодезические даты?
- 4) Какие типы систем координат различаются в ГИС?
- 5) Приведите характеристику Мировой геодезической системы координат WGS-84.
- 6) Приведите характеристики систем криволинейных координат.
- 7) Что такое Проецированные системы координат и картографические проекции?
- 8) Приведите характеристику классификации проекций по типу искажений.

- 9) Приведите характеристику классификации проекций по типу поверхности проецирования.
- 10) Какие параметры определяют картографическую проекцию?
- 11) В чем состоит отличие универсальной трансверсальной проекции Меркатора от проекции Гаусса-Крюгера.
- 12) Опишите методы географических преобразований.

Часть 4

**ОРГАНИЗАЦИЯ ДАННЫХ
В ГЕОИНФОРМАЦИОННЫХ
СИСТЕМАХ**

Раздел 4.1

ПОНЯТИЕ СИСТЕМНОЙ ОРГАНИЗАЦИИ ДАННЫХ

4.1.1 Определение, значение и задачи системной организация данных

Организацией данных называют процесс сведения разнородных данных и моделей в единую логически непротиворечивую модель, которую в дальнейшем можно будет эффективно применять в различных технологиях анализа и управления [38].

Собранные данные могут храниться в виде наборов или файлов. Кроме того, при сборе данные могут организовывать связанные совокупности, называемые моделями. Для того чтобы разнородные данные и модели можно было обрабатывать в одной системе они должны быть упорядочены и сведены к единой информационной модели, в которой они будут дополнять друг друга. Результатом организации данных является создание такой информационной модели, которая позволяет организовать эффективное хранение в базе данных и эффективную обработку в информационных системах и различных технологиях. Организация данных придает качественно новые свойства гео данным. Именно организация данных дает возможность использовать географические данные при решении широкого круга прикладных задач управления, анализа, логистики, планирования, проектирования, прогнозирования, использования ресурсов, мониторинга и др.

Исходная и предварительно обработанная информация включает множество параметров, некоторые из которых могут дублировать друг друга. Уменьшение числа данных о реальных объектах достигается применением моделей, сохраняющих основные свойства объектов исследования и не содержащих второстепенных свойств. Одной из особенностей сбора данных в геоинформатике является то, что исходные данные могут иметь не только разные размерности, но и измеряться в разных шкалах измерений. Организация данных в геоинформатике создает условия для сведения данных разных размерностей и шкал измерений в единую среду и их совместного анализа. Именно это создает возможность комплексного анализа данных при работе с разнородными данными, измеренными в разных

шкалах измерений.

В процессе организации данных все многообразие входной информации - об объектах, их характеристиках, о формах и связях между объектами, различные описательные сведения - преобразуются в наборы моделей. При обработке в геоинформационных технологиях используют информационные модели. Интегрированная информационная модель включает совокупность более простых моделей. Для эффективной обработки данных эта совокупность должна быть оптимизирована. Эта задача решается при организации данных.

Выбор того или иного способа организации данных в ГИС, в первую очередь именно той или иной модели данных, имеет ключевое значение. Выбор модели данных напрямую определяет многие функциональные возможности создаваемой ГИС - некоторые функции просто невозможно реализовать для определенных типов организации данных, или они обеспечиваются очень сложными манипуляциями. Организация данных в ГИС напрямую определяет и применимость тех или иных технологий ввода данных. В той же степени от нее зависит достижимая пространственная точность представления графической части информации, возможность получения качественного картографического материала и организации контроля качества карт. В значительной степени способ организации данных в ГИС определяет также достижимое быстродействие системы, например, при выполнении запроса или визуализации на экране. Возможность работать с большими объемами данных или с точными данными по большим территориям тоже связана со способами и формами организации данных. Удобство редактирования и обновления данных, возможности организации многопользовательской работы в режиме редактирования, создания распределенных по сети баз данных - это все тоже связано в первую очередь с организацией данных, и уже во вторую - с конкретным программным обеспечением [27].

На основании изложенного, можно выделить следующие задачи системной организации данных:

- 1) преобразование информации как описательных сведений в модели;
- 2) сведение множества пространственных данных к единой интегрированной информационной модели;
- 3) классификация исходных данных и моделей при преобразовании их в интегрированную модель;
- 4) идентификация данных в процессе преобразования данных в интегрированную модель, чем сохраняется их индивидуальность;

- 5) установление дополнительных связей между геоданными на основе их интеграции;
- 6) унификация исходных данных и создание возможности обработки и анализа данных, измеренных в разных шкалах и с разными размерностями, в единой системе;
- 7) создание базы для решения основной задачи геоинформатики - установления пространственных отношений между пространственными процессами, объектами, явлениями и их характеристиками.

4.1.2 Уровни организации данных в ГИС

Организацию пространственных данных в ГИС определяют особенности информации об объектах реального мира, с которыми мы имеем дело. Информация о реальном мире частично относится к индивидуализированным объектам, частично к некоторому непрерывно распределенному в пространстве свойству, то есть пространственному полю каких-либо характеристик. Организация пространственных данных в ГИС зависит также от того, как разные лица и организации воспринимают реальный мир. Ощущение структуры (умственной модели) в реальном мире зависит от образования, ответственности, цели организации, где человек работает, от того, что некоторые части мира имеют большую значимость по сравнению с другими частями.

Уровни организации данных соотносятся по степени абстракции в виде последовательности, показанной на Рис. 4.1.1.

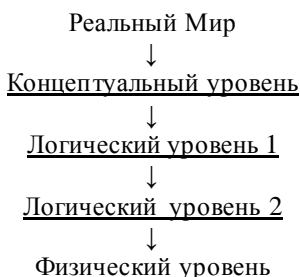


Рис. 4.1.1 – Последовательность уровней организации данных

Концептуальный уровень

Организация данных начинается с трансляции представления пользователя (аналитика) о структуре реального Мира в модели, которые могут быть выполнены на компьютере. Для этого осуществляются следующие мероприятия:

- обследование предметной области, изучение ее информационной структуры;
- выявление всех фрагментов, каждый из которых характеризуется пользовательским представлением, информационными объектами и связями между ними, процессами над информационными объектами;
- моделирование и интеграция всех представлений.

При этом определяют: какие географические объекты представляют область изучения, какие отношения между ними существуют, в каких процессах они участвуют.

На этом уровне организации данных получают *концептуальную модель данных*, инвариантную к структуре базы данных. Часто она представляется в виде модели "сущность-связь".

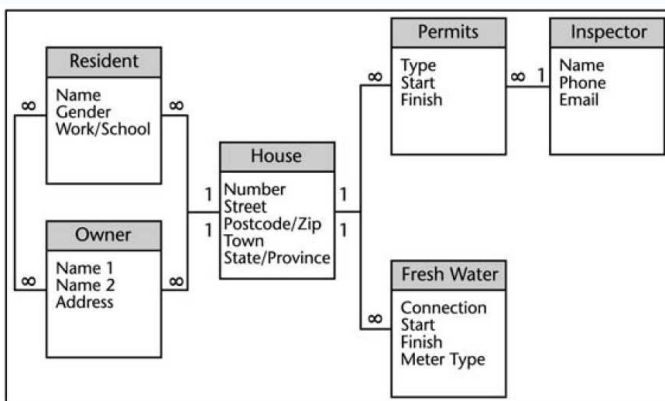


Рис. 4.1.2 – Пример диаграммы "сущность-связь".

Модели данных описывают пространственные сущности и атрибуты, их взаимоотношения друг с другом, как они используются людьми, процессы, которые используются для управления ими, и как данные перемещаются во время использования. Модель данных включает:

- пространственные сущности, которые являются компьютерным представлением пространственных объектов от восприятия

структуры (например, земельный участок, лес, участок трубы, нефтяная скважина),

- описательные данные пространственных сущностей, обычно называемые атрибутами.

Взаимоотношения сущностей выражаются связями (Relationships). Различают классы связей и экземпляры связей. Классы связей – это взаимоотношения между классами сущностей, а экземпляры связи – взаимоотношения между экземплярами сущностей. Класс связей может затрагивать несколько классов сущностей. Число классов сущностей, участвующих в связи, называется степенью связи $n = 2, 3, \dots$. При $n = 2$ связь называется бинарной.

В зависимости от того, сколько экземпляров сущности одного класса связаны со сколькими экземплярами сущности другого класса, различают следующие типы связей:

- связь 1:1 - единичный экземпляр сущности одного класса связан с единичным экземпляром сущности другого класса.
- связь 1:M - единичный экземпляр сущности одного класса связан со многими экземплярами сущности другого класса.
- связь M:N - несколько экземпляров сущности одного класса связаны с несколькими экземплярами сущности другого класса.

Числа, описывающие типы бинарных связей 1:1, 1:M, M:N, обозначают максимальное количество сущностей на каждой стороне связи. Эти числа называются максимальными кардинальными числами, а соответствующая пара чисел называется максимальной кардинальностью.

К этому уровню относятся термины типа полигон, полилиния, дуга, идентификатор, таблица, слой, тема, способ индексирования. Это уровень больше пользователя и администратора базы данных, а также разработчика систем. Все это касается не столько программирования, сколько некоторых разделов математики, с одной стороны, и соотношения элементов модели данных с действительными объектами реального мира, с другой. Основной решаемый на этом уровне вопрос - это вопрос степени адекватности модели данных решаемой задаче. Достаточно ли точно, с сохранением всех ли нужных связей выбранная модель данных позволит смоделировать в компьютере реальную ситуацию?

Логический уровень 1

На более детальном логическом уровне 1 организации данных выполняется преобразование требований к данным в *структуры данных* безотносительно к конкретной СУБД. Для представления структур данных используются записи, элементы данных, связи между записями. Здесь уже фигурируют и математические и программистские термины, такие, как матрицы, списки, системы ссылок, указатели, механизмы сжатия информации. Основные вопросы, которые рассматриваются на этом уровне, это вопросы эффективности в смысле экономии места и быстродействия, то есть уже чисто внутренние компьютерные вопросы эффективности. Логический уровень 1 – это представление программиста.

Логический уровень 2

На следующем по детальности логическом уровне 2 организации данных мы уже имеем дело со *структурой файлов* и их *форматами* применительно к конкретной СУБД. Это размещения информации в файле - что вынесено в его заголовок, и как он организован, что в какой последовательности и какими формами представления чисел записано в этих файлах. Это уровень уже чисто программистский. Рядовому конечному пользователю полезно знать назначение конкретных файлов в определенных структурах данных, возможности и ограничения по использования конкретных форматов файлов.

Физический уровень

На физическом уровне организации данных создается *структура конкретной базы данных* ГИС. Здесь определяются особенности хранения данных, методов доступа, какие объекты вносятся в базу данных, как они распределены между слоями, какие используются классификаторы и т.д. Для описания структуры используются группировки данных, индексы, методы доступа. Уровень организации конкретной структуры базы данных ГИС уникальнй для каждого конкретного проекта.

На основании изложенного, основные характеристики уровней организации данных в ГИС могут быть представлены таблицей 4.1.1.

Табл. 4.1.1 – Характеристики уровней организации данных

Уровень	Информационная модель	Компоненты модели
Концептуальный уровень	Модель данных	Сущности, атрибуты, связи
Логический уровень 1	Структура данных	Записи, элементы данных, связи между записями
Логический уровень 2	Структура файлов, форматы	Конкретные файлы, форматы
Физический уровень	Структура конкретной базы данных	Группировки данных, индексы, методы доступа

Инструментарий

Приведенным уровням моделирования при разработке базирующейся на использовании СУБД прикладной системы соответствует достаточно обширный языковый инструментарий, позволяющий организовать работу с соответствующими моделями. Пример применения современных средств при разработке моделей данных в АО "СПАЭРО Плюс":

- Логические модели предметной области и структуры геобаз данных разрабатываются с помощью CaseTools, поставляемых ESRI в составе ArcGIS. Логическая модель строится в виде набора UML-диаграмм с помощью программного пакета Visio™. При этом используются модифицированные шаблоны ESRI. UML-модель является с одной стороны средством документирования и наглядного представления будущей базы геоданных, а с другой стороны – исходным материалом для описания структуры базы геоданных на языке XML.
- Физическая структура базы геоданных генерируется из XML описания с помощью инструментов ArcCatalog и, при необходимости, дорабатывается с их же использованием.
- Преимуществами применяемой технологии является: полное документирование разработки; наглядность представления структуры базы геоданных, включая все объектные и пространственные классы, их наборы и классы отношений между ними; возможность легко и быстро модифицировать структуру

базы геоданных с полным документированием вносимых изменений.

4.1.3 Принципы организации данных в ГИС

Существует множество структур данных, используемых для представления географических объектов, в зависимости от потребностей в данных и их последующего использования. Наибольшее распространение получили два общепринятых принципа организации моделей пространственных данных [27]:

- принцип послойной организации информации (его часто называют классическим),
- объектно-ориентированный принцип организации данных.

Послойный принцип организации информации

Реальный мир состоит из множества географических объектов. Декомпозиция реального мира может быть выполнена путем стратификации - процедуры, которая разбивает множество географических объектов на страты или слои, удобные для обработки и анализа. В результате изучаемая область реального мира представляется набором слоев взаимосвязанных данных.

Послойный принцип организации информации заключается в том, что данные о территории организуются (расслаиваются) в виде набора тематических слоев. Слой состоит из однородных данных, объединяемых общей тематикой. Например, в один слой выносятся все объекты гидрографии, или все шоссейные дороги, или все, относящиеся к растительному покрову (Рис.4.1.3).

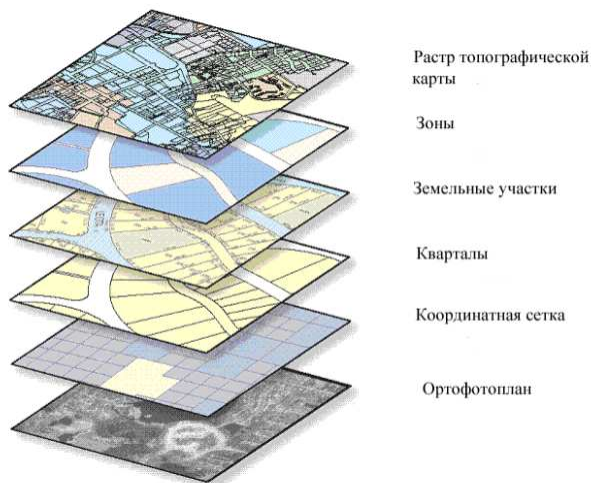


Рис. 4.1.3 - Иллюстрация концепции вертикального расслоения объектов земной поверхности.

Послойный принцип организации информации очень нагляден и хорошо соотносится с приемами традиционной картографии.

Объектно-ориентированный принцип организации данных

Другой принцип появился относительно недавно и связан с объектно-ориентированным подходом, типовым для современного программирования. При этом группировка объектов соответствует их логическим взаимосвязям. Объектно-ориентированный принцип организации данных в ГИС фокусирует внимание не столько на общих свойствах объектов (моделируемых через деление на слои в предыдущем подходе), сколько на их положении в какой-либо сложной иерархической схеме классификации и на взаимоотношениях между объектами. В силу этого удобно отображаются различные родственные и генетические отношения между объектами, отношения соподчиненности, функциональные связи между объектами.

Такой подход ближе к структуре человеческого мышления. Он эффективен, когда необходимо использование логических взаимосвязей объектов, но мало полезен при непрерывном распределении в пространстве признаков (рельеф, удельное содержание полезного ископаемого, загрязнение почвы тяжелыми металлами).

Комбинации принципов организации данных

При комбинации принципов организации данных используются и слои, и отдельные объекты, каждый из которых представляется как отдельный слой. Слои могут быть интегрированы.

4.1.4 Виды моделей организации данных

Принципы организации данных определяют соответствующие модели организации данных:

- 1) геореляционная модель организации данных;
- 2) объектно-ориентированная модель организации данных;
- 3) объектно-реляционная модель организации данных.

Широкое распространение получили модели данных ESRI: а) геореляционная модель организации данных в виде модели данных "Шейпфайл" и модели данных "Покрытие", б) объектно-ориентированная модель организации данных в виде модели данных "База геоданных". На рис. 4.1.4 представлена иерархия моделей данных ESRI от общего верхнего уровня моделей географических объектов до нижнего специального уровня организации данных.

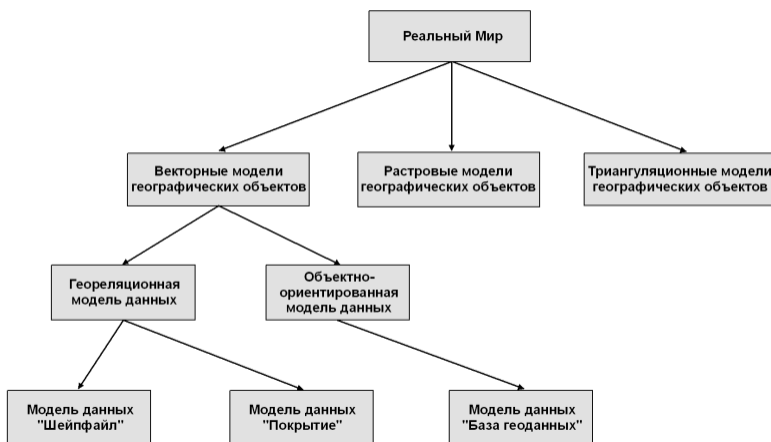


Рис. 4.1.4 – Модели данных ESRI [47]

На рис. 4.1.4 модель данных "База геоданных" представлена как ветвь векторной объектно-ориентированной модели, которая в ней играет ведущую роль. Фактически база геоданных содержит модели данных "шейпфайл" и "покрытие", растровые и триангуляционные модели данных.

4.1.5 Контрольные вопросы и задания для самостоятельной работы

- 1) Что понимается под системной организацией данных?
- 2) Приведите характеристику уровней организации данных.
- 3) Какие принципы организации моделей пространственных данных получили наибольшее распространение?
- 4) Представьте иерархию моделей организации данных.

Раздел 4.2

ГЕОРЕЛЯЦИОННАЯ МОДЕЛЬ ДАННЫХ

4.2.1 Сущность геореляционной модели данных

В базе данных, спроектированной как реляционная модель данных, данные хранятся как наборы таблиц (называемых отношениями), которые логически ассоциированы друг с другом с помощью общих атрибутов. Отдельные записи хранятся как строки таблиц, в то время как атрибуты хранятся в виде колонок. Каждая колонка может содержать атрибутивные данные только одного типа: дату, текстовую строку, числовые данные и т.п. Таблицы обычно стандартизируются для минимизации дублирования.

ГИС содержит два типа данных – пространственные и семантические.

- Пространственные данные географических объектов хранятся в отдельных таблицах пространственных данных в виде последовательности координатных пар X,Y.
- Атрибутивные данные географических объектов организовываются в таблицы атрибутивных данных. Число записей в таблицах атрибутов равно числу графических объектов в двоичных файлах.

Отношения между географическими объектами делаются явными с помощью топологии, которая также представляется соответствующими таблицами.

В ГИС, ориентированных на работу с базами данных (в частности, в ARC/INFO) [28], успешно применяется *геореляционная модель данных* (Georelational Data Model). Сущность этой модели заключается в раздельном хранении значений координат и атрибутивных данных. Эта модель основана на геометрическом типе объекта и отображает мир в виде наборов точек, линий и полигонов. Координаты каждого объекта с уникальным идентификационным номером, хранятся в двоичных "arc"-файлах. Атрибутивные значения и описание топологии хранятся в таблицах реляционной СУБД (изначально в INFO таблице). Записи связаны с геометрией посредством идентификационного номера объекта (Identifier – ID). Модель географических данных представляет географические объекты как набор взаимосвязанных пространственных и атрибутивных данных. При этом ГИС осуществляет совместное согласованное управление целостной

информацией объектов, распределяемой между файловой системой и базой данных.

Таким образом, геореляционная модель данных определяется следующими условиями:

- 1) между записями в таблицах пространственных данных, которые отображают модели географических объектов (точками, линиями, полигонами), и записями в таблице атрибутов устанавливается отношение "один-к-одному";
- 2) связь между географическим объектом и записью в таблице атрибутов поддерживается через единственный уникальный номер – идентификатор объекта;
- 3) идентификатор хранится в двух местах: в файлах географических объектов, содержащих пары координат X,Y, и в соответствующих записях таблицы атрибутов географических объектов.

Важным требованием этой модели является так называемая планарность исполнения, что означает, что плоская поверхность покрывается непрерывно без каких-либо петель неперекрывающихся объектов.

Сила реляционных таблиц в том, что они упрощают реальный мир и дают быстрые и достоверные ответы на запросы, которые они позволяют обрабатывать.

В геореляционных моделях данных используется фиксированный набор встроенных типов данных, таких как числа, даты и текстовые строки.

Геореляционные модели требуют сложного прикладного программирования для результативного моделирования комплексных ситуаций реального мира. Сложные переменные требуют большого количества взаимосвязанных таблиц.

Однако геореляционная модель не способна моделировать все разнообразие географических объектов для конкретной области пользователя. В модели объектов есть геометрические структуры, в которых один и тот же объект (например, полигон) может отображать дом или земельный участок, если обе формы совпадают. Единственная разница заключается в их атрибутах.

4.2.2 Слои пространственных данных - вертикальная организация данных

В большинстве программных продуктов ГИС данные организованы в тематические слои данных. Такой подход позволяет вводить данные как отдельные темы и накладывать их на основе

анализа потребностей. Это может быть представлено концепцией вертикального расслоения объектов земной поверхности. Понимания слоев пространственных данных как вертикальной формы организации данных представлено в пособии [48]. Инструментом пространственной взаимосвязи слоев по вертикали является единая для всех слоев система координат для определения пространственных объектов.

Идея наложения настолько естественна для специалистов картографов и природных ресурсов, что она также была создана при разработке большинства векторных систем САД. Подход наложение/слой, применяемый в системах САД, использован для отделения основных классов пространственных объектов. Эта концепция также используется для логического упорядочения данных в большинстве ГИС программ.

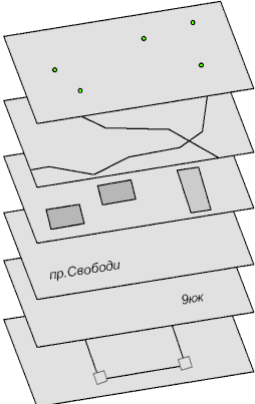
Терминология может отличаться между программным обеспечением ГИС, но подход является тем же. В коммерческом программном обеспечении ГИС используется разнообразие терминов для определения данных слоев. Слой данных (Layer) и Тема (Theme) являются наиболее распространенными терминами; они не являются собственностью какого-либо конкретного программного обеспечения ГИС.

Слой данных или тема соответствует логической совокупности пространственных объектов с общими характеристиками. Каждый слой/тема определяется следующими условиями:

- объекты одного класса,
- одинаковый набор полей.

Геометрический тип объекта определяет 6 классов слоев/тем (Табл.4.2.1).

Таблица 4.2.1 – Классы слоев

Классы слоев	Графическое представление классов слоев
<ol style="list-style-type: none"> 1) Точечный слой (Point Layer) 2) Линейный слой (Line Layer) 3) Полигональный слой (Polygon Layer) 4) Слой аннотаций (Annotation Layer) 5) Слой записей (Record Layer) 6) Слой реперных точек (Tic Layer) 	

В следующей табл.4.2.2 приведены примеры некоторых слоев данных из типичной ГИС муниципального управления.

Таблица 4.2.2 – Примеры слоев данных

<i>Название слоя</i>	<i>Объект реального мира</i>	<i>Геометрический тип объекта</i>	<i>Атрибуты пользователя</i>
Оси улиц	Городские улицы	Линейный	Название, класс улицы
Дорога	Проезжие части улиц	Полигональный	№, площадь, ширина
Кварталы	Кварталы города	Полигональный	Плотность застройки, Численность населения
Здания жилые	Здания	Полигональный	Этажность, количество квартир, количество жильцов,
Здания промышленные	Здания	Полигональный	Высота, объем,

Продолжение табл. 4.2.2.

Земельные участки	Зонирование	Полигональный	Код участка, площадь, тип землепользования
Железная дорога	Основные железно-дорожные линии	Линейный	Название железной дороги
Линии газоснабжения	Система газоснабжения	Линейный	Диаметр, давление, длина
Линии водоснабжения	Система водоснабжения	Линейный	Диаметр, материал, длина
Колодцы водоснабжения	Система водоснабжения	Точечный	Отметка верха, глубина, запорное устройство
Жалобы на канализацию	Местоположения аварийных участков канализации	Точечный	Адрес, дата, код типа аварии

Для любого ГИС проекта потребуются различные слои данных. Они должны быть определены до начала проекта. Также должен быть определен приоритет ввода или оцифровки пространственных данных слоев. Это является обязательным, так как часто один слой данных содержит пространственные объекты, которые совпадают с другими, например, озера могут быть использованы для определения полигонов в слое данных лесного кадастра. Слои, как правило, определяются разработчиком, исходя из потребностей пользователей и доступности данных. Определение слоев может значительно отличаться в зависимости от потребности в ГИС. Количество слоев в действующей ГИС крупного муниципального образования может измеряться несколькими сотнями.

Как правило, в ГИС вводится один слой данных за один раз. Слой данных будет полностью загружен, когда будут выполнены графические преобразования, редактирования, топологические построения, преобразования атрибутов, связывание, а также проверки, прежде чем следующий слой данных будет запущен. Поэтому есть несколько этапов в полной загрузке слоя данных.

Большинство ГИС проектов интегрируют слои данных, для создания производных тем или слоев, которые представляют собой

результат некоторого вычисления или географической модели, например, ценности лесов, пригодности использования земель и т.д. Производные слои полностью зависят от цели проекта. Каждый слой данных, интегрированный в индивидуальном порядке и топологически, будет исходным для создания комбинированных данных слоев. Отдельные функции анализа данных могут быть осуществлены, основываясь на модели данных, например, векторной или растровой, и топологической структуре. Важно отметить, что в векторной ГИС топологическая структура определяется только с помощью уникальных меток для каждого пространственного объекта.

При разработке логической модели данных с использованием реляционной модели может потребоваться рассмотреть послынную и мозаичную структуру данных, известную также как библиотека карт.

Чаще всего при этом организуется также и деление одного тематического слоя по горизонтали по аналогии с отдельными листами карт. Это делается в основном из удобства администрирования баз данных и чтобы избежать работы с чрезмерно большими файлами.

Число слоев при послынной организации данных может быть ограничено, может быть практически не ограничено в зависимости от конкретной реализации. При послынной организации данных очень удобно манипулировать большими группами объектов, представленных слоями, как единым целым, например, включая или выключая слои для визуализации, определять операции, основанные на взаимодействии слоев. В целом можно сказать, что послынная организация данных имеет большой аналитический потенциал.

Объекты, отнесенные к одному слою образуют некоторую физически отдельную единицу данных. Они собираются в один файл или в одну директорию, они имеют единую и отдельную от других слоев систему идентификаторов, к ним можно обращаться как к некоторому множеству. По одной теме может быть предусмотрено несколько слоев разного масштаба и соответственно разной точности или разных временных интервалов. Эта идея также используется для логического упорядочения данных в большинстве ГИС программ.

4.2.3 Пространственная индексация - горизонтальная организация данных

Внутренняя организация слоев данных в горизонтальной форме в ГИС известна как пространственное индексирование [48]. В векторных ГИС пространственные индексы используются для более быстрого доступа к объектам на определенном участке карты. Индексирование пространственных объектов позволяет уменьшить вычислительную сложность процедур поиска пересекающихся и вложенных объектов, поэтому индексы являются важной частью алгоритмов наложения полигонов.

Пространственное индексирование – это метод использования программного обеспечения для хранения и получения пространственных данных. Существуют различные подходы для ускорения процесса поиска пространственных объектов в рамках программного продукта ГИС. Большинство из них связаны с разделением географической области на управляемые подмножества или блоки. Эти блоки затем индексируются математически, например, методом разбиения квадродерева или построения R-дерева, чтобы обеспечить быстрый поиск и извлечение при запросе по инициативе пользователя. Пространственное индексирование аналогично определению листов карты, за исключением того, что конкретные методы индексирования используются для доступа к данным через карту границ листов (блоков). Это делается просто для повышения производительности запросов для больших наборов данных, которые охватывают несколько листов карты, а также для обеспечения целостности данных посредством границ листов карты.

Метод и процесс пространственного индексирования, как правило, прозрачен для пользователя. Тем не менее, он становится очень важным, когда используются особенно большие объемы данных. Понятие пространственной индексации играет все более важную роль в разработке крупномасштабных приложений с использованием ГИС технологий.

Важно отметить, что растровые системы по характеру своей структуры данных обычно не требуют метода пространственного индексирования. Растровый подход налагает регулярные, легко адресуемые разделы по данным Мира неразрывно с ее структурой данных. Таким образом, пространственная индексация, как правило, не требуется. Однако в более сложных векторных ГИС требуется способ быстрого получения пространственных объектов.

4.2.4 Модель данных "Шейпфайл"

Модель данных "Шейпфайл" представляется цифровым форматом Shapefile.

Формат Шейпфайл (Shapefile) – это цифровой векторный формат ESRI для хранения пространственной и связанной семантической/атрибутивной информации о географических объектах. Этот формат не приспособлен для хранения топологической информации. Формат Shapefiles создан для ArcView GIS; он может использоваться в ARC/INFO, ArcGIS.

Формат Shapefile содержит набор файлов с одинаковым названием, но с разным расширением. Например:

1. Rivers.shp
2. Rivers.dbf
3. Rivers.shx
4. Rivers.sbn
5. Rivers.sbx
6. Rivers.ain
7. Rivers.aih

Эти файлы делятся на обязательные и факультативные (дополнительные).

Обязательными файлами являются три файла с расширением .shp, .shx, .dbf, так как они содержат базовые данные.

- *Файл формы* с расширением **.shp (shape file)** – это главный файл, который хранит географические объекты в его собственной записи как список координатных пар **x,y**. Так как каждый географический объект (кроме точечных) может содержать различное число координатных пар, каждая запись в .shp файле может быть различной длины. Записи переменной длины требуют долгое время, чтобы обрабатывать и представлять на экране, поэтому .shp файл индексируется с .shx файлом.
- *Файл индекса формы* с расширением **.shx (shape index file)** содержит одну запись фиксированной длины для каждой записи в .shp файле. Каждая запись в .shx файле содержит номер записи и длину в байтах соответствующей записи в .shp файле. Так как все записи в .shx файле равной длины, они могут читаться быстро и обрабатываться так, как табличный поиск для отчетов в .shp файле. Файл индекс формы (.shx) ускоряет вычерчивание всех пространственных объектов в шейпфайле. Можно ускорить запрос и отбор индивидуальных пространственных объектов, создавая пространственный и атрибутивный индексы.

- *Файл атрибутов* с расширением **.dbf (dBASE file)** хранит атрибутивную информацию о пространственных объектах в .shp файле как таблицу атрибутов в формате dBASE. Таблица атрибутов содержит одну запись для каждого пространственного объекта в .shp файле. Каждая запись атрибутов имеет отношение один-к-одному с соответствующей записью формы.

Дополнительные 2 файла пространственных индексов ArcView GIS создает всякий раз, когда Вы выполняете пространственное соединение или выбор темы темой.

- *Файл пространственного лотка* с расширением **.sbn (spatial bin)** разделяет область, содержащую географические объекты .shp файла в прямоугольные области, названные лотками. Каждый лоток содержит запись числа пространственных объектов в .shp файле, которые попадают в его область. Когда Вы делаете пространственный запрос, записи в .sbn файле читаются первыми, и рассматриваются только пространственные объекты, которые пересекают лотки, указанные запросом. Это значительно ускоряет обработку. Так как различное число пространственных объектов может попадать внутрь каждого лотка, однако, записи лотка изменяются в длине и нуждаются в их собственном индексе.
- *Файл пространственного индекса лотка* с расширением **.sbx (spatial bin index)** файл содержит одну запись фиксированной длины для каждой записи в .sbn файле. Каждая запись в .sbx файле содержит запись числа и длины в байтах соответствующей записи лотка в .sbn файле. Так как записи в .sbx все одинаковой длины, они могут читаться быстро и обрабатываться так, как табличный поиск в записях переменной длины в .sbn.

Дополнительные 2 файла индекса атрибута ArcView GIS создает всякий раз, когда Вы выполняете связь с таблицей.

- *Файл индекса атрибута* с расширением **.ain (attribute index)** файл содержит один индекс для каждого поля, включенного в связь, или такого, для которого создан индекс методом, описанным выше. Эти индексы улучшают операции с данными типа соединения и связи, и ускоряют простые запросы типа: [Название] = "Харьков". Индекс атрибута не будет улучшать операции на запросах, которые включают строки соответствия типа [Название], или сравнение [Количество] < 100000.
- *Файл заголовка индекса атрибута* с расширением **.aih (attribute index header)** файл содержит название каждого поля, которое было индексировано. Он служит как каталог к значениям, содержащимся в .ain файле.

К дополнительным файлам относятся еще три файла.

- *Файл системы координат* с расширением **.prj** (map **pr**ojection) хранит информацию о системе координат и картографической проекции.
- *Файл метаданных* с расширением **.xml** (extensible **marku**p language) хранит данные о данных в ArcInfo и ArcView 8.0 и выше для использования в Internet.
- *Файл легенды* с расширением **.avl** (Arc**V**iew **leg**end) хранит символику графического отображения объектов.

Шейпфайлы являются простыми, поскольку они хранят примитивные геометрические типы данных точечные, линейные и полигональные. Эти примитивы имеют ограниченное использование без каких-либо признаков для указания того, что они представляют. Таким образом, таблица записей будет хранить пространственные объекты / атрибуты для каждой примитивной формы в шейпфайле. Формы (точечные, линейные, полигональные), а также данные атрибутов могут создавать бесконечное множество представлений о географических данных. Представление предоставляет возможности для мощного и точного вычисления.

4.2.5 Модель данных "Покрытие"

Покрытие (Coverage) - это геореляционная модель, которая имеет векторный топологический формат данных. Институтом ESRI разработаны две версии файлов покрытия: для PC ARC/INFO в 1981 г., для версий ArcGIS с 2000 г.

В математике Покрытие множества X – есть семейство подмножеств этого множества, объединение которых есть X , или семейство подмножеств пространства, в котором расположено X и которое содержит X . Элементы покрытия несут в себе полную информацию о локальном строении пространства. Обычно понятие покрытия рассматривается в контексте общей топологии.

Покрытие содержит пространственные и атрибутивные данные географических объектов. Покрытие использует набор классов пространственных объектов для представления географических объектов.

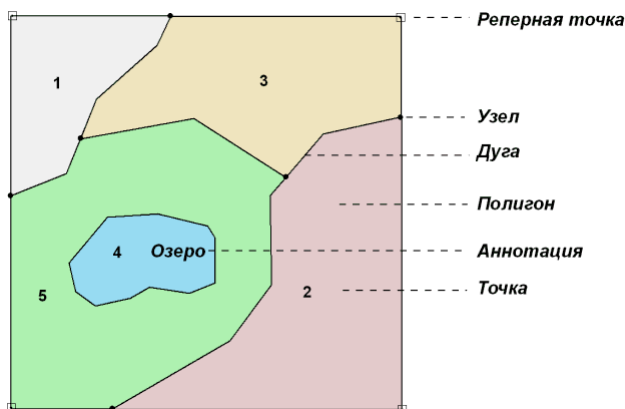


Рис. 4.2.1 - Классы пространственных объектов в покрытии

Модель данных Покрытие использует следующие классы пространственных объектов:

- *Точка (Point)* – используется для представления точечных пространственных объектов или пользовательских идентификаторов ID полигонов. Точка определяется координатной парой x, y .
- *Дуга (Arc)* – используется для представления линейных пространственных объектов или границ полигонов. Дуга определяется последовательностью координатных пар x, y начального узла, промежуточных вершин, конечного узла. Дуги топологически связываются через их конечные точки (узлы). Один линейный объект может быть образован многими дугами.
- *Узел (Node)* – представляет конечные точки дуг или пересечение линейных объектов. Узел имеет уникальный идентификатор. Узел может быть топологически связан с набором дуг, которые соединены одна с другой.
- *Путь (Route)* – линейный пространственный объект, состоящий из одну или несколько дуг или части дуг.
- *Секция (Section)* – дуга или часть дуги, которая используется для определения пути или создания путевых блоков.
- *Полигон (Polygon)* - представляет площадные объекты. Полигоны топологически определяются серией дуг, которые формируют их границы, включая дуги, определяющие острова внутри. Пользовательские идентификаторы ID полигонов представляются точками внутри границ.

- *Регион (Region)* – совокупность полигонов, представляющих географический объект.
- *Аннотация (Annotation)* – текст, используемый для обозначения объектов. Аннотации не имеют топологических связей с другими объектами, не используются для аналитических целей.
- *Реперная точка (Tic)* – регистрационная точка, определяющая положение известной точки на земной поверхности, для которой известны координаты реального земного пространства. Реперные точки позволяют регистрировать и трансформировать координаты покрытия. Количество реперных точек рекомендуется 4 и более.
- *Охват покрытия (Coverage Extent)* – минимальный прямоугольник, ограничивающий покрытие, который представляет территориальный охват покрытия. Охват покрытия определяется предельными координатами X_{max} , X_{min} , Y_{max} , Y_{min} его элементов.

Описательные данные для классов пространственных объектов хранятся в соответствующих таблицах атрибутов. Связывание пространственных объектов и атрибутов обеспечивается следующими базовыми положениями:

- Пространственные объекты в покрытии существуют в отношении один-к-одному с соответствующими записями в таблице атрибутов пространственных объектов;
- ArcGIS поддерживает связь между пространственными объектами и атрибутами посредством уникального идентификатора, назначенного каждому объекту;
- Порядковый номер пространственного объекта физически хранится в двух местах покрытия: в файлах, содержащих пространственные данные для каждого пространственного объекта (координатные пары) и с соответствующей записью в таблице атрибутов пространственных объектов. ArcGIS автоматически создает и поддерживает эти связи.

Набор классов в покрытии варьирует в зависимости от географических объектов, которое оно представляет. Покрытие может содержать:

- 1) набор точек, которые представляют географические объекты, и ассоциированные таблицы атрибутов, которые описывают эти точечные объекты, или
- 2) набор узлов и дуг, которые представляют линейные пространственные объекты, и ассоциированные таблицы атрибутов, которые описывают эти линейные объекты, или

- 3) набор узлов и дуг, которые окружают географические области (полигоны), и ассоциированные таблицы атрибутов, которые описывают эти области, или
- 4) комбинацию наборов 1) и 2), или
- 5) комбинацию наборов 2) и 3).

Отсюда следует, что дуга и узел являются главными "строительными" элементами покрытия. Ключевое значение дуги Arc используется в названии программных продуктов ArcInfo, ArcGIS.

Покрытие может также содержать другие элементы, такие как аннотации и реперные точки.

Покрытие хранится как директория, в которой каждый класс пространственных объектов хранится как набор файлов. Например, покрытие "дороги" есть линейное покрытие, содержащее файл дуг (Arc), файл аннотаций (Annotation) для дуг, файл реперных точек (Tic). Пространственные данные хранятся в двоичных файлах, а атрибутивные и топологические данные хранятся в таблицах INFO. Покрытие также может иметь ассоциированные файлы.

Топология имеет отношение к способу, согласно которому линейные данные хранятся и соотносятся. Топология регистрирует пространственные отношения между дугами и полигонами в покрытии. Каждая дуга имеет "от-узла", "к-узлу", полигон слева и справа, уникальный идентификатор ID дуги и внутренний порядковый номер. Группы дуг, которые формируют замкнутые формы (полигоны), ассоциируются с уникальной меткой. Хранение данных таким образом позволяет системе определять, какие полигоны смежные, какие дуги формируют полигоны, как далеко отстоят центры дуг и полигонов один от другого и так далее. Следует иметь в виду, что топология не создается автоматически.

Пользователь использует специальные команды для создания топологии.

4.2.6 Преимущества и недостатки геореляционной модели данных

Геореляционная модель имеет следующие преимущества:

1. простая структура таблиц, которые легко читать;
2. интуитивный, простой пользовательский интерфейс;
3. наличие множества инструментов для конечных пользователей (например, макросов и скриптов);

4. простота изменения и добавления новых привязок, данных и записей;
5. простота использования таблиц, описывающих географические элементы с общими атрибутами;
6. возможность привязки таблиц, описывающих топологию, необходимую для ГИС-анализа;
7. прямой доступ к данным, обеспечивающий их быструю и эффективную обработку;
8. независимость данных от приложения;
9. наличие больших объемов ГИС-данных в этом формате.

Геореляционная модель имеет следующие недостатки:

1. ограниченное представление реального мира;
2. ограниченную гибкость управления запросами и данными;
3. медленный последовательный доступ;
4. трудность моделирования сложных отношений данных, поскольку для этого часто необходимы квалифицированные прикладные программисты баз данных;
5. необходимость выражения сложных отношений в виде процедур в каждой программе, которая обращается к базе данных.

4.2.7 Контрольные вопросы и задания для самостоятельной работы

- 1) В чем состоит сущность геореляционной модели данных?
- 2) Что представляют собой тематические слои и какие принципы необходимо соблюдать для их организации ?
- 3) Для чего выполняется пространственное индексирование?
- 4) Представьте содержание модели данных "Шейпфайл".
- 5) Опишите содержание модели данных "Покрытие".
- 6) В чем состоят преимущества и недостатки геореляционной модели данных?

Раздел 4.3

ОБЪЕКТНО-ОРИЕНТИРОВАННАЯ МОДЕЛЬ ДАНЫХ

4.3.1 Основные положения объектно-ориентированной методологии

Воспринимаемая сложность реального мира и ее познание требуют развития соответствующих методов и средств, в том числе геоинформационных систем. В свою очередь, развитие геоинформационных систем связано с необходимостью совместной обработки увеличивающихся объемов пространственной и непространственной информации, разработки более сложных процессов обработки взаимосвязанной разноплановой информации, интеграции этой информации во взаимодействии с другими различными по назначению системами. Дополнительные требования в части нахождения лучших решений, удобства, производительности, надежности и стоимости также требуют разработки и развития адекватных моделей.

Распространение мощных персональных компьютеров создало в 90-х годах основу для широкого применения объектно-ориентированного подхода в практике проектирования и программирования информационных систем. Новая методология ориентирована, прежде всего, на преодоление сложности, связанной с разработкой программных средств, на создание больших сложных систем, коллективную их разработку, последующее активное сопровождение при эксплуатации и регулярные модификации [9].

4.3.1.1 Понятие объекта

В основе объектно-ориентированной методологии лежит объектный подход, при котором предметная прикладная область представляется в виде совокупности объектов, которые взаимодействуют между собой посредством передачи сообщений [49].

Под *объектом* понимается некоторая сущность (реальная или абстрактная), обладающая состоянием, поведением и индивидуальностью.

- *Состояние объекта* характеризуется перечнем всех его возможных свойств - структурой и значениями каждого из этих свойств.
- *Поведение объекта* (или его функциональность) характеризует то, как объект взаимодействует с другими объектами или подвергается взаимодействию других объектов, проявляя свою индивидуальность. Поведение объекта реализуется в виде функций, которые называют методами. При этом структура объекта доступна только через его методы, которые в совокупности формируют интерфейс объекта.
- *Индивидуальность объекта* характеризуют такие свойства объекта, которые отличают его от всех других объектов.

Сами по себе объекты не представляют никакого интереса: только в процессе взаимодействия объектов реализуется система. Для объектно-ориентированной методологии представляют особый интерес два типа иерархических соотношений объектов:

- *связи* - обозначают равноправные отношения между объектами; объект сотрудничает с другими объектами через связи, соединяющие его с ними;
- *агрегация* - агрегация описывает отношения целого и части, приводящие к соответствующей иерархии объектов

Объекты образуют минимальные единицы инкапсуляции.

Инкапсуляцией называется способ объединения структуры и поведения в одном месте (как бы в "капсуле") и сокрытия всех данных внутри объекта, что делает их невидимыми для всех, за исключением методов самого объекта. В объектно-ориентированной модели объекты инкапсулируют атрибуты и линии поведения.

Доступ к данным, заключенным в объекте, возможен только в соответствии с линиями поведения объекта. Таким образом, инкапсуляция защищает данные от повреждения другими объектами, а также закрывает внутренние детали объектов от остальной системы. Инкапсуляция также обеспечивает степень независимости данных, чтобы не возникало необходимости изменять объекты-отправители или получатели сообщений при их взаимодействии с объектом, поведение которого изменилось. Инкапсуляция — это сущность объектно-ориентированной модели.

4.3.1.2 Понятие класса

Одно из ключевых понятий объектно-ориентированного подхода – понятие класса. Под *классом* понимается множество объектов,

имеющих общую структуру и общее поведение. Именно класс вначале описывает переменные и методы объекта, то есть структуру и поведение объекта. Любой конкретный объект представляет собой экземпляр класса. Объекты, не связанные общностью структуры и поведения, нельзя объединить в класс, так как по определению они не связаны между собой ничем.

Существенными являются следующие *типы отношений* между классами:

- отношение "обобщение/специализация" (общее и частное) отражает степень общности,
- отношение "целое/часть" отражает агрегирование объектов;
- отношение "ассоциация" отражает смысловую связь между классами, которые не связаны никакими другими типами отношений;
- наследование - это такое отношение между классами, когда один класс повторяет структуру и поведение другого класса (одиночное наследование) или других (множественное наследование) классов. Класс, структура и поведение которого наследуются, называется суперклассом. Производный от суперкласса класс называется подклассом. Это означает, что наследование устанавливает между классами иерархию общего и частного.

4.3.1.3 Понятие отношения

Отношения описывают то, как объекты ассоциированы друг с другом. Они определяют правила создания, изменения и удаления объектов. Существует несколько видов отношений, которые могут использоваться в объектно-ориентированной модели данных.

- *Наследование* — позволяет одному классу наследовать атрибуты и линии поведения одного или нескольких других классов. Класс, наследующий атрибуты и линии поведения, известен как подкласс. Родительский класс называется суперклассом. Помимо наследуемой ими линии поведения, подклассы могут добавлять или переопределять унаследованные атрибуты и линии поведения. Суперкласс — это генерализация его подклассов, а подкласс — это уточнение своего суперкласса. Например, дом — это уточнение здания, а здание — это генерализация дома. Класс домов может наследовать атрибуты и линии поведения класса зданий, такие как количество этажей, комнат и тип сооружения.

- *Ассоциация* — общие отношения между объектами. Каждая ассоциация может также обладать ассоциированной с ней множественностью, которая определяет количество объектов, ассоциированных с другим объектом. Например, ассоциация может сказать вам, что "владелец" объекта может владеть одним или многими домами. Объединение (агрегирование) и композиция — это особые типы ассоциаций.
- *Объединение (агрегирование)* — определенный тип ассоциации. Объекты могут содержать другие объекты, поэтому объединение — это просто набор разных классов объектов, собранных в один класс, который становится новым объектом. Классы объектов могут быть собраны в объединенный класс. Например, объединенный класс "Земельно-имущественный комплекс" может быть создан путем объединения классов "Земельный участок" и "Здания". Эти новые составные объекты важны, потому что они, в отличие от простых объектов, способны представлять более сложные структуры.
- *Композиция* — еще одна специальная форма ассоциации. Это — более сильная ассоциативная взаимосвязь, при которой жизнь "содержимых" классов объектов управляет жизнью "содержащего" класса объектов (контейнера). Например, здание составлено из основания, стен и крыши. Если вы удалите здание, то автоматически удалите его основание, стены и крышу, но не его символ.

4.3.1.4 Принципы объектной модели

Объектно-ориентированная технология основывается на так называемой объектной модели. Основными принципами, на которых строятся объектные модели, являются: абстрагирование, инкапсуляция, модульность, иерархичность, типизация, параллелизм и сохраняемость.

- *Абстракция* выделяет существенные характеристики некоторого объекта, отличающие его от всех других видов объектов и, таким образом, четко определяет его концептуальные границы с точки зрения наблюдателя.
- *Инкапсуляция* - это процесс отделения друг от друга элементов объекта, определяющих его устройство и поведение; инкапсуляция служит для того, чтобы изолировать контрактные обязательства абстракции от их реализации.

- *Модульность* - это свойство системы, которая была разложена на внутренне связанные, но слабо связанные между собой модули.
- *Иерархия* - это ранжирование абстракций, расположение их по уровням.
- *Типизация* - это способ защититься от использования объектов одного класса вместо другого, или, по крайней мере, управлять таким использованием.
- *Параллелизм* - это свойство, отличающее активные объекты от пассивных.
- *Сохраняемость* - способность объекта существовать во времени, переживая породивший его процесс, и (или) в пространстве, перемещаясь из своего первоначального адресного пространства.

4.3.2 Общая характеристика объектно-ориентированной модели данных "База геоданных"

4.3.2.1 Определение базы геоданных

С 2000 года в ArcGIS представлен новый подход к хранению и представлению географических данных - объектно-ориентированная модель данных, названная "базой геоданных" - БГД (Geographic Database - GDB),

База геоданных – это совокупность географических наборов данных различных типов, используемых в ArcGIS, содержащихся в общих папках файловой системы или в реляционной базе данных.

Модель данных ГИС База геоданных базируется на принципах реляционных таблиц. Модель данных использует персональную базу данных Microsoft Access или многопользовательскую реляционную базу данных, такую как Oracle, Microsoft SQL Server, PostgreSQL, Informix, или IBM DB2.

Ключевой концепцией базы геоданных является набор данных. База геоданных содержит три основных типа наборов данных:

- классы пространственных объектов (Feature classes);
- растровые наборы данных (Raster datasets);
- непространственные таблицы (Tables).

Эти наборы данных добавляют в базу геоданных с расширенными возможностями (например, путем создания топологии сети или подтипов) для моделирования поведения, сохранения целостности

данных и работы с одним из важных наборов пространственных отношений.

База геоданных обеспечивает:

- доступ и управление большими объемами географических данных, хранящимися в файлах и базе данных,
- обработку богатых и разнообразных типов данных и других объектов,
- применение сложных правил и отношений в "интеллектуальных" ГИС, непосредственно моделирующих реальность.

4.3.2.2 Объектно-ориентированная векторная модель данных

Модель базы геоданных поддерживает объектно-ориентированную векторную модель данных. В этой модели сущности представлены как объекты со свойствами, поведением и отношениями. Поддержка различных типов географических объектов встроена в систему. К этим типам объектов относятся простые объекты, географические объекты, сетевые пространственные объекты, аннотации пространственных объектов и другие, более специализированные типы пространственных объектов. Модель позволяет определить взаимоотношения между объектами, а также правила для поддержания целостности ссылочных данных между объектами.

Объектно-ориентированная модель данных рассматривает географические объекты реального мира как объекты базы данных. Объектами представляются пространственные объекты, например, дорога, здание, земельный участок. Объекты могут представлять слой дорог, систему координат слоя дорог.

Модель базы геоданных определяет общую модель для географической информации. Это типовая модель, которая может быть использована для определения и работы широким кругом различных пользователей или приложений по конкретным моделям.

4.3.2.3 Преимущества базы геоданных

Преимущества базы геоданных в предоставлении возможности:

- централизованно хранить географические данные и управлять ими в одной реляционной СУБД;
- моделировать поведение пространственных объектов;

- применить сложные правила и отношения к данным;
- поддерживать целостность пространственных данных в непротиворечивой, точной базе данных;
- работы в рамках многопользовательского доступа и редактирования среды;
- масштабирования созданных решений;
- интеграции пространственных данных с другими базами данных;
- поддержки пользовательских функций и поведения.

4.3.3 Средства придания интеллектуальных свойств пространственным объектам

В модели данных покрытия пользователь может описывать только геометрию объекта и его характеристику, где описание его поведения возможно только с помощью дополнительных созданных пользователем программ, что ограничивает возможности анализа и обработки, например анализ разных типов объектов в реальном времени и т.д.

Модель данных базы геоданных обладает ключевым преимуществом – возможностью построить интеллектуальную модель пространственной системы [50]. С помощью модели данных базы геоданных пользователь может создавать более содержательные объекты с новыми качествами (интеллектуальные объекты) и, тем самым, моделировать объекты реального мира.

Пользователь работает не просто с обычными точками, линиями и полигонами, информация о которых хранится в таблицах. В БГД пользователь может оперировать такими понятиями, как объекты реального мира, устанавливать и настраивать свойства и взаимоотношения объектов. Например, вместо точек можно работать с трансформаторами, а вместо линий - с трубами. При этом каждой трубе будет установлено правило, через какой переходник она соединяется с другой трубой.

Пользователь может задавать поведение отдельных объектов, определять взаимоотношения классов объектов, создавать правила и применять топологические модели высокого уровня без программирования.

Модель данных База геоданных имеет средства создания более содержательных пространственных объектов, которыми моделируется поведение, поддерживается целостность данных и работа с пространственными отношениями. Такими средствами являются,

прежде всего, топология, подтипы классов пространственных и непространственных объектов, домены атрибутов, отношения, правила связности и другие.

4.3.3.1 Топология

Определение и назначение топологии

В базе геоданных *топология* - это механизм, который определяет, как точечные, линейные, полигональные пространственные объекты используют общую геометрию. Например, оси улиц и почтовых участков имеют общую геометрию, а смежные полигоны почв имеют общие границы.

Топология определяет и обеспечивает правила целостности данных (например, не должно быть щелей между полигонами). Она поддерживает запросы топологических отношений и перемещений (например, перемещение пространственных объектов смежности или связности), поддерживает сложные инструменты редактирования, а также позволяет создавать пространственные объекты из неструктурированных геометрий (например, создание полигонов из линий).

В ArcGIS топология включает следующие компоненты:

- 1) топологические модели данных, использующие открытый формат для простых пространственных объектов, правила топологии и топологически интегрированные координаты для пространственных объектов с общей геометрией;
- 2) топологические слои, которые используются для отображения топологических взаимоотношений, ошибок и исключений;
- 3) комплекс инструментов геообработки для создания, анализа, управления и проверки топологии;
- 4) передовые программные алгоритмы для анализа и нахождения топологических элементов в классах пространственных объектов точечных, линейных и полигональных;
- 5) богатые средства редактирования и автоматизации данных, которые используются для создания, поддержания и проверки топологической целостности и представления общих пространственных объектов.

Свойства топологии в базе геоданных

В базе геоданных для каждой топологии определяются следующие свойства:

- имя топологии, которая будет создана;
- кластерный допуск, определяющий минимальное допустимое расстояние между вершинами объектов;
- список классов пространственных объектов, которые примут участие в топологии;
- ранги относительной точности координат для каждого класса пространственных объектов;
- список правил топологии.

Правила топологии

Топология реализуется в виде набора правил целостности, определяющих поведение пространственно взаимосвязанных географических объектов и объектных классов. Эти правила позволяют моделировать в базе геоданных такие пространственные отношения, как связность (связаны ли линии улично-дорожной сети?) и смежность (существует ли промежуток между двумя полигонами участков?). Правила топологии определяют допустимые пространственные отношения между пространственными объектами. Правила используются для контроля топологических отношений между пространственными объектами внутри одного класса пространственных объектов, между пространственными объектами в различных классах или между подтипами пространственных объектов.

Таблица 4.3.1- Примеры применения правил.

Правило	Пример применения
Не должны перекрываться	Земельные участки не должны перекрываться; смежные участки могут иметь общие границы. Это правило может использоваться для управления целостностью пространственных объектов.
Должны совмещаться с объектами класса	Территория большого города должна быть полностью покрыта площадными объектами его административных районов.
Граница должна совпадать с	Границы планировочных кварталов должна совпадать с красными линиями застройки.
Должны совпадать с	Линии автобусных маршрутов должны совпадать с осевыми линиями дорог.
Не должны пересекаться	Горизонталы не должны пересекаться.

В ArcGIS формализовано 25 правил топологии.

4.3.3.2 Подтипы

Подтипы (Subtypes)- это подмножество пространственных объектов в классе пространственных объектов или объекты в таблице, которые имеют одинаковые атрибуты.

Подтипы используются для создания расширенных возможностей базы геоданных.

- 1) Подтипы позволяют *повысить эффективность базы геоданных*.
 - Подтипы используются в качестве метода для классификации данных, для разделения классов пространственных объектов или таблиц на логические группы, базирующиеся на значениях атрибутов. Например, улицы в классе пространственных объектов улиц можно было бы разбить на три подтипа: улицы местного значения, улицы районного значения и общегородские улицы. Представление различных объектов реального мира как подмножество пространственных объектов в данном классе пространственных объектов заменяет создание нового класса пространственных объектов для каждого объекта.
- 2) Подтипы позволяют *управлять значениями атрибутов*, в том числе:
 - Установить значение по умолчанию, которое будет автоматически применяться при создании новых пространственных объектов. Например, подтип улиц местного значения может быть создан и определен так, что когда этот подтип улицы будет добавлен классу пространственных объектов улиц, его атрибут максимальной скорости будет автоматически установлен на 40 км. в час.
 - Применить домен диапазонов или домен кодированных значений к пространственным объектам, чтобы ограничить ввод информации согласно установленному набору значений. Например, в водопроводной сети, подтип магистральных водопроводов может иметь домен кодированных значений для строительных материалов, ограничивающий их выполнение из железа, чугуна или меди.
- 3) Подтипы позволяют *расширить правила поведения*, в том числе:
 - Создать правила соединения между другими подтипами и классами пространственных объектов для сохранения

целостности сети. Например, в водопроводной сети концы труб разного диаметра могут быть соединены специальным переходным элементом.

- Создать правила топологии между другими подтипами и классами пространственных объектов, находящимися в топологии. Например, можно внести требование о том, что пространственные объекты улиц должны быть связаны с другими пространственными объектами улиц с обеих сторон, за исключением подтипов тупиковых улиц.
- Разработать правила отношений между другими подтипами, таблицами и классами пространственных объектов. Например, в электрической сети можно создать правила отношений между подтипами, которые заключаются в том, что стальные столбы поддерживают класс трансформаторов В, а деревянные столбы поддерживают класс трансформаторов С.
- Создать пользовательские правила между пространственными объектами с помощью написания кода.

Подтипы могут быть созданы, используя контекстное меню в ArcCatalog или используя набор инструментов подтипов (Subtypes) геообработки.

4.3.3.3 Домены

Два превосходных механизма Домены и Подтипы используются для организации данных так, что операции управления, отображения и редактирования данными становятся более эффективными при поддержке целостности атрибутов.

Домен – это декларация (фиксированный список) допустимых значений атрибута. Всякий раз, когда домен ассоциируется с полем атрибутов, только значения из домена будут правильными для этого поля. Другими словами, поле не примет значение не из домена.

В базе геоданных домен атрибутов является механизмом для усиления целостности данных. Домены атрибутов определяют, какие значения являются допустимыми в поле класса пространственных объектов или таблице непространственных атрибутов. Если пространственные или непространственные объекты сгруппированы в подтипы, различные домены атрибутов могут ассоциироваться с каждым подтипом.

База геоданных использует два типа доменов атрибутов.

- *Домены диапазонов* создаются на интервальных данных. Домен диапазонов задает допустимый диапазон значений для числового атрибута. При создании ряда доменов необходимо ввести минимальное и максимальное значение. Например, для подтипа линий газоснабжения низкого давления задается диапазон диаметров труб от 40 мм. до 270 мм.
- *Домены кодов* создаются на номинальных данных. Домен кодированных значений может применяться к любому типу атрибутов – текстовых, цифровых, дат и так далее. Домен кодированных значений определяет правильный набор значений для атрибута. Например, утвержденный городским советом стандарт на названия улиц, являющийся доменом кодов, включает, в том числе значение атрибута "Чернышевская ул."; поле названий улиц не примет значение атрибута "Чернышевского ул."

4.3.3.4 Отношения и классы отношений

Различные виды географических и негеографических сущностей могут быть связаны между собой отношениями. Отношения могут существовать между:

- географическими сущностями и другими географическими сущностями, например, здание может быть связано с земельным участком;
- географическими сущностями и негеографическими сущностями, например, земельный участок может быть связан с владельцем;
- негеографическими сущностями и другими негеографическими сущностями, например, собственник земельного участка может быть связан с налоговым кодом.

При представлении связей между географическими объектами необходимо моделировать пространственные отношения между пространственными объектами.

ArcMap обеспечивает два метода ассоциации данных, хранящихся в таблицах с географическими объектами: связывание и соединение.

Связывание (Relating) - операция, которая устанавливает временные связи между записями в двух таблицах, используя общий ключ в обеих. Связывание реализуют типы отношений "один-к-одному" (one-to-one), один-ко-многим" (one-to-many) и "многие-ко-многим" (many-to-many).

Соединение (Joining) – операция связывания и физического соединения двух таблиц атрибутов, используя их общие элементы,

общих для двух таблиц. Соединение обычно используется для того, чтобы дополнить поля одной таблицы к полям другой таблицы посредством общих атрибутов или полей. Соединение реализуют типы отношений "один-к-одному" (one-to-one) и "многие-к-одному" (many-to-one).

Отношения один-к-одному и один-ко-многим не требуют новой таблицы в базе геоданных. Связывание и соединение используются для создания данных, изучения и анализа.

В базе геоданных отношения между объектами хранятся в классах отношений. *Класс отношений* (Relationship class) – это элемент базы геоданных, который хранит информацию об отношениях. Классы отношений управляют ассоциациями между объектами в одном классе (классе пространственных объектов или таблице) и объектами в другом классе.

Классы отношений обеспечивают множество *продвинутых возможностей*, не имеющихся в операциях связывания и соединения в ArcMap.

1) Классы отношений помогают обеспечивать *целостность ссылочных данных*.

- Класс отношений может быть настроен так, что при изменении объекта связанные объекты обновляются автоматически. Это может включать физическое перемещение связанных пространственных объектов, удаление связанных объектов или обновление атрибута. Например, можно создать такие отношения, при которых при перемещении опоры электросети перемещаются вместе с ней прикрепленные трансформаторы и другое оборудование.
- Устанавливая правила, класс отношений может ограничить тип отношений. Например, одна опора может поддерживать не более трех трансформаторов. Стальная опора может поддерживать трансформаторы класса В, но не трансформаторы класса С.

2) Классы отношений *облегчают редактирование*, помогая снизить расходы на техническое обслуживание.

- Обеспечивая автоматическое обновление для связанных объектов, класс отношений может избавить от выполнения дополнительных операций редактирования.
- Классы отношений облегчают вам *доступ к объектам*, когда вы редактируете. Вы можете выбрать один объект, а затем найти все связанные объекты, используя диалоговое окно

атрибутов или таблицы. Выйдя на соответствующий объект, вы можете редактировать его атрибуты.

- 3) Классы отношений позволяют делать запросы к связанным пространственным объектам и записям, запрашивать, выполнять анализ и формировать отчеты с атрибутами из класса отношений.



Рис. 4.3.1 – Класс отношений "один-ко-многим" в базе геоданных

Класс отношений может иметь атрибуты. Класс отношений, который имеет атрибуты, хранится в таблице базы геоданных. Каждое отношение хранится как запись в таблице класса отношений (Рис.).

Классы отношений имеют следующие *характеристики*: имя, атрибуты отношения, класс-источник и класс-адресат, первичные и внешние ключи, тип отношения, мощность, направление сообщения информации.

- Все отношения в классе отношений связывают объекты из одного класса-источника (origin class) с объектами из одного класса-адресата (destination class).
- Таблица класса отношений имеет пару внешних ключей (foreign key - FK), один из которых относится к классу-источнику и другой – к классу-адресату. Классы объектов имеют внутренние ключи (primary key - PK), один из которых относится к классу-источнику и другой – к классу-адресату.
- Класс отношений имеет метку прямого направления (forward path label) и метку обратного направления (backward path label). Примеры меток направления: "управляет", "управляется посредством".
- Мощность отношения (cardinalities) определяет количество объектов в классе-источнике, которые могут относиться к числу объектов в классе-адресате. Отношения в базе геоданных могут иметь одну из трех значений мощности: "один-к-одному", "один-ко-многим", или "многие-ко-многим".

- База геоданных поддерживает два типа отношений - простые и составные.
Простое отношение (simple relationship) – это равноправное отношение, при котором связанные объекты могут существовать независимо друг от друга.
Составное отношение (composite relationship) – это отношение один-ко-многим, при котором объекты из класса-адресата не могут существовать независимо от объектов из класса-источника. Когда источник удаляется, соответствующие объекты из класса-адресата также уничтожаются.

4.3.4 Элементы БГД

База геоданных содержит три основных типа наборов данных независимо от системы их использования [50]:

- таблицы (Tables);
- классы пространственных объектов (Feature classes);
- наборы растровых данных (Raster datasets).

Эти наборы данных можно рассматривать как универсальную отправную точку для разработки базы геоданных.

База геоданных имеет ряд дополнительных элементов и типов наборов данных, которые могут быть использованы для моделирования сложного поведения, повышения целостности данных, для расширения возможностей базы геоданных в управлении данными (Рис. 4.3.2).

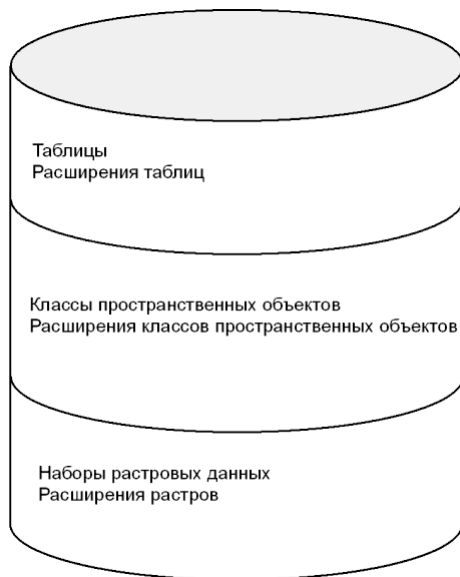


Рис. 4.3.2 - Компоненты базы геоданных

Все три основных типа наборов данных в базе геоданных, а также другие элементы базы геоданных хранятся с использованием таблиц.

4.3.4.1 Таблицы

В базе геоданных атрибуты управляются в таблицах, базирующихся на серии простых существенных концептах реляционных данных:

- таблицы содержат ряды;
- все ряды таблицы имеют одинаковые колонки;
- каждая колонка имеет тип данных;
- серия реляционных функций и операторов доступна для операций на таблицах и их элементами данных.

Таблицы и отношения играют ключевую роль в ArcGIS, так как они находятся в приложениях традиционных баз данных. Пользователи выполняют многие традиционные табличные и реляционные операции, используя таблицы.

Таблицы (Tables) обеспечивают описательной информацией географические объекты, растры и традиционные таблицы атрибутов в

базе геоданных. Ряды в таблице могут использоваться для хранения всех свойств географических объектов. Они включают хранение и управление геометрией пространственных объектов в колонке Shape.

4.3.4.2 Расширения таблиц

Для моделирования отношений и поведения таблицы могут использоваться совместно с элементами, которые добавляют продвинутое возможности базе геоданных. База геоданных содержит следующие элементы, которые образуют *расширения таблиц* (Extending tables).

- *Домены атрибутов* (Attribute Domains) – представляют списки допустимых значений или диапазона допустимых значений атрибутов столбцов. Домены используются для обеспечения целостности значений атрибутов и классификации данных.
- *Классы отношений* (Relationship Classes) – создают отношения между двумя таблицами, используя общий ключ. Классы отношений определяют строки во второй таблице, соответствующие выбранным строкам в первой таблице
- *Подтипы* (Subtypes) - управляют наборами атрибутов подклассов в одной таблице. Подтипы часто используются в таблицах классов пространственных объектов для управления различными поведением в подмножествах того же типа пространственных объектов.
- *Управление версиями* (Versioning) – управляют длинными транзакциями обновления данных, историческими архивами и многопользовательским редактированием.

4.3.4.3 Классы пространственных объектов

Классы пространственных объектов (Feature Classes) – это гомогенные собрания пространственных объектов с одинаковым пространственным представлением, с общей системой координат и набором атрибутов, хранящихся в таблице базы геоданных.

В базе геоданных основные классы пространственных объектов имеют следующие 7 типов:

- 1) *Точки* (Points) — используются для представления географических объектов, размерами которых можно пренебречь для данной задачи. Точками характеризуют местоположение географических объектов.

- 2) *Линии (Lines)* — используются для представления географических объектов, которые имеют длину, но не имеют площади. Линиями представляют форму и местоположение географических объектов.
- 3) *Полигоны (Polygons)* — набор многосторонних площадных пространственных объектов, которые представляют форму и местоположение однородных типов пространственных объектов.
- 4) *Аннотация (Annotation)* — текст карты, содержащий свойства представления текста.
- 5) *Размеры (Dimensions)* — специальный тип аннотации, который показывает определенную длину или расстояние. Например, длина стороны здания или расстояние между точками.
- 6) *Мультиточки (Multipoints)* — используются для представления пространственных объектов, которые образованы более чем одной точкой с общими атрибутами.
- 7) *Мультифрагменты (Multipatches)* — используются для представления внешней поверхности или оболочки пространственных объектов, которые занимают отдельную область или объем в трехмерном пространстве. Мультифрагменты включают плановые 3D кольца и треугольники, которые используются в комбинации для моделирования трехмерной оболочки. Мультифрагменты могут использоваться для представления любых объектов от простых, таких как сферы или кубы, до составных объектов, таких как изоповерхности или здания.

4.3.4.4 Расширения классов пространственных объектов

Для моделирования пространственных отношений и поведения классы пространственных объектов могут использоваться совместно с элементами, которые добавляют продвинутое возможности базе геоданных. База геоданных содержит следующие элементы, которые образуют расширения классов пространственных объектов (Extending feature classes).

- 1) *Набор классов пространственных объектов (Feature dataset)* – собрание классов пространственных объектов, которые имеют общую систему координат.
- 2) *Подтипы (Subtypes)* - используются с таблицами класса пространственных объектов для управления различным поведением в подмножествах одного типа пространственных объектов. Подтипы управляют набором подклассов

пространственных объектов в одном классе пространственных объектов.

- 3) *Домены атрибутов* (Attribute Domains) - определяют список допустимых значений или диапазоны допустимых значений атрибутов столбцов. Домены используются для обеспечения целостности значений атрибутов и для классификации данных.
- 4) *Классы отношений* (Relationship Classes) - создают отношения между классами пространственных объектов и другими таблицами с использованием общего ключа.
- 5) *Топология* (Topology) – моделирует общую геометрию пространственных объектов.
- 6) *Сетевой набор данных* (Network Dataset) – моделирует транспортную связность и потоки.
- 7) *Геометрические сети* (Geometric Network) – моделирует коммунальные сети и трассирование.
- 8) *Набор данных местности* (Terrain Dataset) – моделирует сеть нерегулярных треугольников (TIN) и управляет большими лидарными и гидролокационными коллекциями точек.
- 9) *Адресный локатор* (Address Locator) – адресное геокодирование.
- 10) *Набор данных кадастрового материала* (Cadastral Fabric Dataset) – интегрирует и поддерживает геодезическую информацию для подразделений и планов земельных участков, как часть модели данных непрерывного кадастрового материала в базе геоданных. Набор данных кадастрового материала вносит дополнительные улучшения точности для материала земельного участка, как только будут введены новое подразделение планов и описание земельного участка.
- 11) *Линейное моделирование* (Linear Referencing) – определяет местоположение событий вдоль линейных пространственных объектов с измерениями.
- 12) *Картографические представления* (Cartographic Representations) – управляет множеством картографических представлений и правилами улучшенного картографического черчения.
- 13) *Управление версиями* (Versioning) – руководит некоторыми ключевыми потоками работ в ГИС для управления данными.

4.3.4.5 Наборы растровых данных

Растровые данные представляют собой географические объекты путем деления территории на дискретные квадратные или

прямоугольные ячейки, вложенные в сетку. Каждая ячейка имеет значение, которое используется для представления некоторых характеристик для этого места.

Растровые данные обычно используются для управления и представления изображений, цифровых моделей рельефа, а также ряда других явлений. Нередко растры используются как способ представления точечных, линейных и полигональных пространственных объектов.

Растры интересны по двум причинам: во-первых, они могут быть использованы для представления всей географической информации (пространственных объектов, изображений и поверхностей), во-вторых, они имеют богатый набор аналитических операторов геопроецирования.

4.3.4.6 Расширения растров

Растры интенсивно и все в большей степени используются в ГИС приложениях. База геоданных может управлять растрами для многих целей: в качестве индивидуальных наборов данных, в качестве логической коллекции наборов данных, а также атрибуты изображений в таблицах.

Растры могут использоваться совместно с элементами, которые добавляют продвинутые возможности базе геоданных. База геоданных содержит следующие элементы, которые образуют *расширения растров* (Extending rasters).

- 1) *Набор растровых данных* (Raster Datasets) – управляет очень большими непрерывными наборами данных изображений и мозаиками изображений.
- 2) *Каталоги растров* (Raster Catalogs) – используются для ряда целей: управлять слоем листов изображений, где каждый лист есть изображение; управлять серией изображений в СУБД; управлять серией временных растров.
- 3) *Колонки атрибутов растров в таблицах* (Raster Attribute Columns in Tables) – хранит изображения или сканированные документы как атрибуты в таблице.
- 4) *Определение сервиса изображения* (Image Service Definition) – сервисы публикации изображений для коллекции растровых данных в поле на диске, в каталоге изображений и базе геоданных.

4.3.5 Типы баз геоданных

Существует два типа баз геоданных - *персональные и многопользовательские* БГД.

В *персональных* базах геоданных данные хранятся *локально* на отдельном компьютере. Поддержка персональных баз геоданных встроена в настольные продукты ArcGIS. Локальная версия базы геоданных может храниться и на настольном компьютере в виде файла Microsoft Access (*.mdb). Использование персональных баз геоданных разумно при малых объемах данных и когда не требуется использовать многопользовательские сеансы доступа и редактирования.

В *многопользовательских* базах геоданных данные хранятся на сервере, используя в качестве хранилища данных известные коммерческие РСУБД, такие как IBM DB2, Informix, Oracle, SQL Server, PostgreSQL. Для этого следует применить ArcSDE - специальное приложение семейства ArcGIS. ArcSDE устанавливается на сервер данных для взаимодействия с РСУБД в рамках корпоративной ГИС. Многопользовательский вариант базы геоданных имеет смысл использовать для крупных организаций, когда объемы данных превышают отметку с шестью нулями,

Пользователь может выбрать в качестве начальной основы локальные версии баз геоданных. По мере роста объема накопленных данных пользователь может постепенно осуществлять переход на многопользовательскую корпоративную РСУБД. Этот переход не потребует кардинальной реорганизации производственного ГИС-процесса, что очень важно. При этом локальная база геоданных сохранит всю структуру данных и наследует правила и свойства всех объектов, заданные в базе геоданных на сервере.

Многопользовательские базы геоданных позволяют редактировать одни и те же географические данные многим пользователям одновременно на нескольких рабочих местах. Например, в организации несколько отделов, и каждый из отделов должен участвовать в рабочем процессе, причем, желательно, а иногда и необходимо, чтобы все эти отделы работали сообща и одновременно - в реальном времени. Для этого требуется правильно организовать рабочий процесс и всегда иметь под рукой удобный инструмент для управления этим процессом. При такой постановке рабочего процесса его участники должны оперировать своего рода версиями (копиями) общей базы данных, с которыми они взаимодействуют.

Версия представляет моментальную копию всей базы геоданных. Она содержит все наборы данных в базе геоданных. Версия изолирует

работу пользователей посредством множества сессий редактирования, позволяя пользователям редактировать без блокировки пространственные объекты в выполняемой версии.

Эффективность описанных выше возможностей, которые обеспечивает многопользовательская база геоданных, очевидна.

4.3.6 Преимущества и недостатки объектно-ориентированной модели данных

Объектно-ориентированная модель базы геоданных имеет ряд *преимуществ*:

1. обеспечивает комплексное представление реального мира;
2. модель интуитивна, поскольку в ней используются объекты, существующие в реальном мире;
3. инкапсуляция, объединяя атрибуты и линии поведения объекта, делает возможным доступ к объекту с помощью четко определенного набора методов и атрибутов без знания содержимое объекта;
4. обеспечивает высокий уровень целостности данных (новые данные должны следовать правилам поведения);
5. обеспечивает моделирование сложных отношений между данными;
6. поддерживает множественные уровни генерализации, объединения и ассоциации;
7. хорошо интегрируется с методами имитационного моделирования;
8. имеет функцию поддержки версий для множественного одновременного обновления;
9. требует меньше кодирования в ГИС-программах, что означает меньше ошибок и более низкую стоимость поддержки.

У объектно-ориентированной модели данных есть также и некоторые *недостатки*:

1. комплексные модели реального мира сложнее разрабатывать и строить;
2. большие и комплексные модели выполняются медленнее;
3. эта модель зависит от тщательности описания явлений реального мира (что особенно трудно в мире природы);
4. анализ объектно-ориентированных баз данных требуют использования объектно-ориентированных языков программирования;
5. некоторые бизнес-приложения могут не иметь доступа к объектно-ориентированной базе данных или возможности передачи в нее данных.

4.3.7 Контрольные вопросы и задания для самостоятельной работы

- 1) Что понимается под объектом в объектно-ориентированной методологии?
- 2) Что понимается под классом в объектно-ориентированной методологии?
- 3) Какие виды отношений используются в объектно-ориентированной модели данных?
- 4) Приведите общую характеристику модели данных ГИС "База геоданных".
- 5) Приведите общую характеристику топологии в модели данных ГИС "База геоданных".
- 6) Каково назначение Подтипов и Доменов модели данных ГИС "База геоданных"?
- 7) Как реализуются отношения и классы отношений в базе геоданных?
- 8) Приведите общую характеристику таблиц и расширений таблиц как компонентов базы геоданных.
- 9) Приведите общую характеристику классов пространственных объектов и расширений классов пространственных объектов как компонентов базы геоданных.
- 10) Приведите общую характеристику растровых данных и расширений растров как компонентов базы геоданных.
- 11) Какие существуют типы баз геоданных?

Часть 5

**ОСНОВЫ
ГЕОПРОСТРАНСТВЕННОГО
АНАЛИЗА**

Раздел 5.1

ОБЩАЯ ХАРАКТЕРИСТИКА ГЕОПРОСТРАНСТВЕННОГО АНАЛИЗА

5.1.1 Определение геопространственного анализа

Геопространственный анализ – это процесс поиска пространственных закономерностей в распределении географических данных и взаимосвязей между объектами [51]. Пространственный анализ – это сердце ГИС [11]. Геопространственный анализ предоставляет особые точки зрения на мир. Геопространственный анализ является уникальной линзой, через которую изучаются события, структуры и процессы, которые происходят на земле или вблизи поверхности нашей планеты [52]. В результате анализа географической информации получается качественно новая информация и выявляются прежде неизвестные закономерности.

Мишель Ф. Гудчайлд в предисловии к книге Енди Митчелла [51] пишет, что процесс пространственного анализа напоминает растяжение резиновой ленты, когда длинная и тяжелая работа по цифровке элементов карт, формированию баз данных, выявлению ошибок и трансформирование информации в разнообразные системы координат, в конце концов, вознаграждает эффективным результатом или находкой оптимального решения.

Областью геопространственного анализа является поверхность Земли, оболочка над ней при анализе топографии и атмосферы, оболочка под ней при анализе грунтовых вод и геологии. *Масштаб* сущностей простирается от небольших объектов (например, записи археологов о местах кусочков керамических изделий размером в несколько сантиметров или границ собственности измеренных до миллиметра) до глобальных (например, анализ температуры поверхности моря и глобального потепления). Анализ простирается в прошлое *время* (в исторические исследования миграции населения, в изучение структуры в археологических местонахождениях или в детальное картирование движения континентов) и в будущее (в попытках предсказать направления ураганов, таяния льдов или рост городских районов). Методы пространственного анализа работают в ряде пространственных и временных масштабов.

В конечном счете, геопространственный анализ имеет отношение к проблеме "что" происходит "где". Геопространственный анализ использует географическую информацию, которая является базовой информацией для построения на ней структур и аргументов, которая обеспечивает богатство пространственного анализа. В принципе нет никаких ограничений на сложность пространственных аналитических методов, которые, возможно, нашли применение в мире, и могут быть использованы для стимулирования интересных идей и поддержки практических действий и решений. В действительности, некоторые методы могут быть более простыми, более полезными и более глубокими, чем другие.

Пространственный анализ существует на стыке между человеком и компьютером. Оба в нем играют важную роль – с одной стороны, путь человеческой интуиции со всей его расплывчатостью и неформальностью, и с другой стороны, путь формального, точного пространственного анализа мира.

5.1.2 Задачи геопространственного анализа

Преимущество геоинформационной методологии состоит в том, что ГИС позволяет идентифицировать, поддерживать и управлять пространственными *связями* между топологическими объектами, представляющими объекты реального мира, создавать новые объекты, связи, связывать новые атрибуты.

Энди Митчелл в руководстве по ГИС-анализу [51] определяет наиболее общие задачи геопространственного анализа, ежедневно выполняемые людьми на их рабочих местах:

- a) анализ местоположения объектов – поиск, где размещаются объекты (Mapping where things are);
- b) анализ распределения числовых показателей – выявление, где больше, где меньше (Mapping the most and least);
- c) построение карт плотности – картографирование плотности (Mapping density);
- d) поиск объектов внутри области – поиск того, что внутри (Finding what's inside);
- e) анализ окружения – поиск того, что рядом (Finding what's nearby);
- f) анализ пространственных изменений – картографирование изменений (Mapping change)

Решение этих вопросов с помощью собственных средств ГИС и привлеченных внешних моделирующих систем дает, например, возможность:

- обосновать местоположение учебного заведения или бизнес-центра с учетом многих, в т.ч. и пространственных факторов;
- прогнозировать развитие и следствия социологических и экономических ситуаций, стихийных бедствий и аварий естественного техногенного характера в пространстве и времени;
- найти оптимальную трассу трубопровода или путепровода, который проектируется;
- определить место оптимального расположения нового месторождения, а также вычислить его экономическую эффективность;
- оценить эффективность деятельности милиции, служб охраны окружающей среды, пожарников и т.п.;

5.1.3 Методология геопространственного анализа

Одним из средств пространственного анализа является картографическое моделирование. Картографическое моделирование – это общая, и в то же время четкая методология, которая используется многими приложениями ГИС явным способом. Этим термином определяется то, что картографическое моделирование включает модели геопространственной информации, представленные в картографической форме, то есть, как карты.

Картографическое моделирование используется для одновременного анализа пространственных и тематических характеристик геопространственной информации. Тематический компонент геопространственной информации анализируется посредством статистических операций с данными (например, получение среднего значения и среднего квадратического отклонения), а пространственные характеристики геопространственной информации получают методами пространственного анализа.

Аналитические методы могут быть как очень простыми – при обычном создании карты, так и более сложными, включающими модели, которые имитируют реальный мир путем объединения многих слоев информации.

В большинстве случаев геопространственный анализ выполняется в следующей последовательности [51, 52]:

Этап 1. Постановка проблемы

Формулирование проблемы всегда начинается с определения цели анализа в виде вопросов, на которые необходимо получить ответы, например:

- где произошло более всего краж со взломом в прошлом месяце?
- сколько леса растёт в пределах каждого бассейна водосбора?
- какая степень риска строительства в обозначенном месте поймы реки?

Специфика проблемы чаще всего определяет и выбор типа анализа; и метод, наиболее эффективный в данном случае; и способ интерпретации результатов. От характера проблемы зависит также необходимая детальность решения, которая в свою очередь определяет затраты на получение данных, приобретение или аренду необходимых программных средств и вычислительных мощностей. Именно на этом этапе формируется соответствие между масштабом возникшей проблемы и средствами, необходимыми для ее решения.

Важным фактором, который в значительной мере определяет детальность исследования и метод решения данной проблемы, есть представления о назначении результатов анализа. В одном случае надо провести предварительные исследования, чтобы оценить правомерность избранного метода или выделить значимые факторы, в другом – предоставить результаты в отчете на ученом совете. В последнем случае методы должны быть более строгими, а результаты – более обоснованными.

На этом этапе определяются также критерии, которые определяют параметры использования базы геоданных для получения ответов. Например, земельный участок для строительства должен иметь градиент не более 10 %.

Этап 2. Оценка исходных данных

Тип данных и объектов, доступных для проведения данного исследования, в значительной мере определяет специфику метода, который будет использован, и достижимую точность результата. С другой стороны, чтобы получить качественную информацию, нужно обеспечить соответствующий уровень исходных данных.

Как правило, база геоданных уже будет существовать. Нужно четко представлять, какой информацией вы располагаете, и что еще нужно получить или создать. Создание новых данных, в свою очередь, может послужить причиной появления новых атрибутов в таблице данных или даже новых слоев карт. На этом этапе может возникнуть необходимость подготовки данных для пространственных операций, в

том числе, изменения данных, преобразования единиц измерения и системы координат, добавления данных, конвертирования данных из одного формата в другой.

Оценка исходных данных – важнейший этап аналитического процесса. Именно в этот момент определяется принципиальная возможность реализации избранных методов анализа и получения результата заявленного качества.

Этап 3. Выбор метода анализа

Почти всегда есть несколько путей получения необходимой информации. При выборе пути следует опираться на следующие положения.

Во–первых, при выборе метода всегда возникает дилемма: оперативность или точность анализа. Оперативность анализа возникает, когда требуется быстро оценить ситуацию в целом и принять решение. При этом пользуются простыми, хорошо проверенными методами, которые не нуждаются в существенных затратах на получение детальной и всесторонней исходной информации о состоянии объекта. Полученный результат будет иметь невысокую точность и отображать только общие характеристики процесса, который изучается. Точность анализа требуется для получения достоверной и полной исходной информации. Точность анализа требует увеличения затрат времени и усилий на обработку данных.

Во–вторых, метод анализа определяется моделью имеющихся данных. Векторные данные являются наиболее удобными, когда необходимо сохранять точное местоположение исходного объекта, работать с дискретными объектами, границами или моделировать линейную сеть. Растровые данные целесообразно использовать для анализа непрерывных явлений. Триангуляционные данные целесообразно использовать для анализа поверхностей. Современные ГИС позволяют интегрировать в процессе анализа эти типы данных. В случае необходимости преобразования можно конвертировать растровые данные в векторные и, наоборот, с помощью вмонтированных функций.

В–третьих, в зависимости от выбранных моделей данных выбираются и средства их обработки. Современные ГИС имеют большое количество аналитических средств пространственных и атрибутивных данных. Пространственный анализ определенного набора данных может включать операции, например, извлечения объектов, построения буферных зон, наложения буферных зон на

другие слои, работу с объектами, попавшими в буферные зоны, и другие операции. Набор средств анализа определяется при интерпретации критериев анализа, выделенных на этапе 1. Каждое утверждение в постановке проблемы может транслироваться в ряд операций анализа.

Этап 4. Обработка данных

Как только выбран метод, необходимо выстроить цепочку его реализации средствами ГИС. Каждая пространственная операция приводит к новой информации. В большинстве случаев анализа требуется набор операций с множеством слоев. При работе с векторными наборами данных они выполняются ступенчатым образом – два входных слоя используются для формирования нового слоя, этот промежуточный слой обрабатывается совместно с третьим слоем, чтобы сформировать другой промежуточный слой, и так далее до достижения желательного результирующего слоя карты. При работе с растровыми наборами данных есть возможность одновременной обработки нескольких слоев, алгоритм которой реализуется в растровом калькуляторе (Raster Calculator).

Наличие модуля построителя моделей процессов (Model Builder) в ArcGIS 9.x позволяет аналитику автоматизировать выполнение алгоритма модели без услуг программиста.

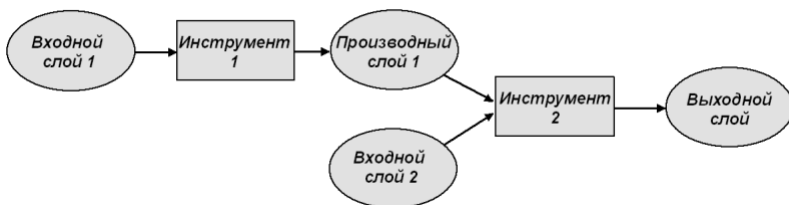


Рис. 5.1.1 - Пример диаграммы процесса анализа в построителе моделей ModelBuilder

Несмотря на довольно большой перечень методов и подходов, которые применяются в современном ГИС-анализе и моделировании, практически всегда с их помощью получаем ответы на два главных вопроса:

1. Где расположенные объекты с заданными свойствами?
2. Почему исследуемые объекты размещены именно в этом месте и в настоящее время?

Этап 5. Оценка и отображение результатов

В процессе оценки результатов выполняется интерпретация результатов, определяется объективность и достаточность полученной информации. Если необходимо, то принимается решение о повторении анализа с другими параметрами, или уточнение анализа, или применение другого метода. ГИС позволяет сравнительно легко и оперативно сделать необходимые изменения и получить новый результат. Можно также оперативно сравнить результаты разных анализов и увидеть, который из подходов оказался лучшим.

Результаты анализа могут быть представлены в виде карты, диаграммы, значений в таблице – фактически новой информации. Необходимо решить, какую информацию выносить на карту, как группировать значения для наилучшего отображения данных.

Эта методология использует как *векторные*, так и *растровые* модели реального земного пространства.

5.1.4 Классификации аналитических средств ГИС

Геоинформационные системы содержат богатый набор разнообразных аналитических средств для проведения операций с географическими объектами. Существует ряд подходов для классификации аналитических средств ГИС. Отсутствие единой их классификации порождает разные трактования и взгляды на суть аналитического процесса вообще.

В ряде фундаментальных работ [53], [54], [55] получила распространение классификация базовых аналитических средств на основе выполняемых функций, предложенная Стеном Ароноффом (Stan Aronoff) [56]. Согласно этой классификации базовые аналитические средства сгруппированы в 4 обширные категории:

1. Функции измерений, выбора данных, классификации;
2. Оверлейные функции;
3. Функции окрестности;
4. Функции связности.

Некоторые функции, не вошедшие в эту классификацию, являются специальными функциями геопроессинга. Следует также иметь в виду, что во многих случаях процесс анализа выполняется посредством использования набора как аналитических, так и специальных операций.

5.1.5 Контрольные вопросы и задания для самостоятельной работы

- 1) Приведите определение и общие задачи геопространственного анализа.
- 2) В какой последовательности выполняется геопространственный анализ?
- 3) Какая классификация базовых аналитических средств получила распространение?

Раздел 5.2

ФУНКЦИИ ИЗМЕРЕНИЙ

Функции измерения (Measurement functions), как и выбора данных и классификации, позволяют анализировать данные без выполнения существенных изменений. Они часто используются в начале анализа. Функции геометрических измерений включают вычисления местоположения, длин линий, расстояний между двумя объектами и площади отдельных объектов.

5.2.1 Измерения на векторных данных

5.2.1.1 Определение местоположения

Свойство "местоположение" пространственных объектов описывается координатами, которые всегда хранятся в базе данных в виде списка координатных пар. В ГИС используются разные инструменты получения координат, в том числе внешне простейшее средство - курсором. Для некоторых аналитических задач представляет интерес определение координат особых точек.

Определение координат точки пересечения двух прямых

Операция нахождения точки пересечения линий является одной из базовых в ГИС-анализе, так как она используется в ряде операций геопроецирования и в оверлейных операциях.

Исходным является уравнение прямой, известное из аналитической геометрии:

$$y = kx + b \quad (5.2.1)$$

где:

b – отрезок, отсекаемый прямой на оси ОУ;

k – угловой коэффициент прямой $k = tg\beta$;

β – угол между положительным направлением оси ОХ и прямой.

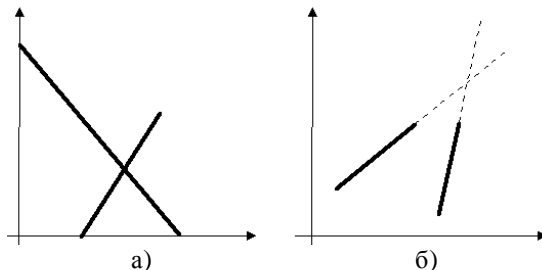


Рис. 5.2.1 – Точка пересечения: а) внутри отрезков прямых; б) снаружи отрезков прямых.

Две линии задаются уравнениями

$$y_1 = k_1 x + b_1 \quad (5.2.2)$$

$$y_2 = k_2 x + b_2 \quad (5.2.3)$$

Совместное решение этих двух уравнений позволяет найти координаты точки пересечения этих двух линий:

$$x = \frac{b_1 - b_2}{k_2 - k_1} \quad (5.2.4)$$

$$y = \frac{k_2 b_1 - k_1 b_2}{k_2 - k_1} \quad (5.2.5)$$

В оверлейных операциях при определении принадлежности точки полигону учитывается специальный случай, когда одна линия параллельна выбранной оси, например, оси OX. В этом случае $k = 0$, а координаты точки пересечения таких двух линий будут равны:

$$x = \frac{b_1 - b_2}{k_2} \quad (5.2.6)$$

$$y = b_1 \quad (5.2.7)$$

Для определения пересечения двух полилиний с n_1 и n_2 сегментами разработано несколько способов.

Самым простым способом нахождения их точек пересечения является последовательная проверка пересечения каждого сегмента первой линии с каждым сегментом второй линии. Сложность этого алгоритма пропорциональна произведению $n_1 * n_2$. Она может быть уменьшена, если предварительно проверять на пересечение экстенты (Extent - минимальный ограничивающий прямоугольник) сегментов и полилиний, представленные на рис.5.2.2.

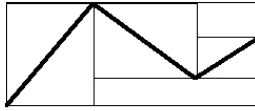


Рис. 5.2.2 - Экстенты сегментов и полилинии

Второй метод, использованный в ГИС ArcInfo, основан на разбиении полилинии на секции, в которых линия монотонно возрастает или убывает по x и по y . Разбиение происходит в точках локального минимума или максимума по x или по y . Горизонтальная или вертикальная линия пересекает такую секцию только в одной точке.

Определения координат центроидов и центров

Термины центры и центроиды имеют различные значения и формулы. Центры и центроиды полигонов выполняют ряд важных функций в ГИС. Они часто используются как "идентификационные точки" полигонов. Для аналитических целей они используются в качестве объектов, которые представляют полигоны. Эффект замещения точками полигонов заключается в меньших объемах данных, а также в возможности выполнять некоторые аналитические операции. Целесообразно использовать центры полигонов, когда полигоны небольших размеров, однородные или относительно компактные. В растровых ГИС представление площадных ячеек раstra их центрами расширяет возможности геопространственного анализа непрерывных поверхностей.

Точечными объектами, которые представляют полигон, могут быть: 1) центр экстента - c_1 , 2) срединный центр - c_2 , 3) центроид - c_3 (Рис.).

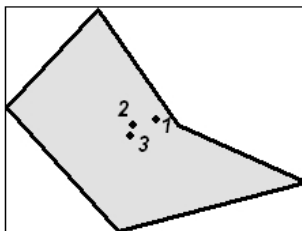


Рис. 5.2.3 - Точки замещения полигонального объекта:
1 - центр экстенга, 2 - срединный центр, 3 – центроид.

В некоторых случаях центры и центроиды могут находиться вне полигона. ArcGIS включает функцию Features to Points, которая может создавать центры внутри полигона.

Центр экстенга полигонального объекта вычисляется по формулам:

$$x_{c1} = \frac{x_{\max} + x_{\min}}{2}; \quad y_{c1} = \frac{y_{\max} + y_{\min}}{2} \quad (5.2.8)$$

Срединный центр (Mean Centre) определяется срединной точкой или точкой симметрии геометрической фигуры. Его местоположение определяется средними координатами:

$$x_{c2} = \sum_{i=1}^{i=n} \frac{x_i}{n}; \quad y_{c2} = \sum_{i=1}^{i=n} \frac{y_i}{n} \quad (5.2.9)$$

Срединный центр не есть центр притяжения. В некоторых случаях используется взвешенный срединный центр (Weighted Mean Centre), который определяется взвешенными средними координатами:

$$x_{c2_p} = \sum_{i=1}^{i=n} \frac{P_i x_i}{\sum_i P_i}; \quad y_{c2_p} = \sum_{i=1}^{i=n} \frac{P_i y_i}{\sum_i P_i} \quad (5.2.10)$$

Центроид (Centroid) рассматривается как центр притяжения объекта или набора объектов. Координаты центроида получают по формулам Гаусса:

$$x_{c3} = \frac{1}{6S} \sum_{i=1}^{n-1} (x_{i+1} y_i - x_i y_{i+1}) * (x_i + x_{i+1}) \quad (5.2.11)$$

$$y_{c3} = \frac{1}{6S} \sum_{i=1}^{n-1} (x_{i+1} y_i - x_i y_{i+1}) * (y_i + y_{i+1}) \quad (5.2.12)$$

где S – площадь полигонального объекта.

В треугольнике центр притяжения находится на пересечении прямых линий, проведенных из вершин до средних точек противоположных сторон.

Определение длин линий

Свойство "длина" характеризует линейные пространственные объекты. Значение длины хранится в базе данных или вычисляется. В декартовой системе координат длина линии зависит от координат двух точек и вычисляется по формуле:

$$d = \sqrt{(x_b - x_a)^2 + (y_b - y_a)^2} \quad (5.2.13)$$

Свойство "расстояние" характеризует удаление одного объекта от другого. Например, пусть прямая линия задана выражением:

$$ax + by + c = 0 \quad (5.2.14)$$

где a, b, c – коэффициенты. Расстояние от точки $P(x_p, y_p)$ до прямой вычисляется по формуле:

$$d_{p,l} = \frac{|ax_p + by_p + c|}{\sqrt{(a^2 + b^2)}} \quad (5.2.15)$$

Определение площади полигона

Для определения площади полигона, заданного последовательностью вершин, чаще всего применяется алгоритм Симпсона (Simpson Thomas), основанный на разбиении многоугольника на трапеции, ограниченные линейными сегментами границы полигона, перпендикулярами, опущенными из вершин сегмента на ось x , и осью x (Рис.). Для сегмента, соединяющего вершины (x_i, y_i) и (x_{i+1}, y_{i+1}) , площадь такой трапеции равна

$$S_{i,i+1} = \frac{1}{2} (x_{i+1} - x_i) * (y_i - y_{i+1}) \quad (5.2.16)$$

Площадь полигона равна сумме площадей трапеций для всех сегментов полигона.

$$S = \frac{1}{2} \sum_{i=1}^{n-1} (x_{i+1} - x_i) * (y_i - y_{i+1}) \quad (5.2.17)$$

Для сегментов, у которых $x_i > x_{i+1}$, площадь получается отрицательной. Следует заметить, что полигон – замкнутая фигура, поэтому нужно учитывать сегмент, соединяющий последнюю вершину с первой.

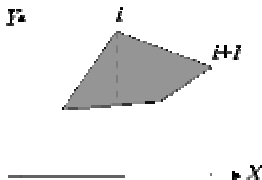


Рис. 5.2.4 - Вычисление площади полигона путем разбиения на трапеции.

Формула вычисления площади полигона может быть преобразована к следующему виду:

$$S = \frac{1}{2} \sum_{i=1}^{n-1} (x_{i+1} y_i - x_i y_{i+1}) \quad (5.2.18)$$

Формула вычисления площади полигона представлена для правой системы координат и нумерации вершин полигона по ходу часовой стрелки. Для полигонов, оцифрованных против часовой стрелки, площадь получается отрицательной. Этим способом можно вычислить площади не только для выпуклых многоугольников, но и для вогнутых, а также для полигонов, имеющих дыры. Алгоритм непригоден для вычисления площадей полигонов, имеющих самопересечения границ.

5.5.2 Измерения на растровых данных

Измерения на растровых данных проще из-за регулярности ячейки. Размер площади ячейки постоянен, и определяется разрешающей способностью ячейки.

Местоположение отдельной ячейки определяется точкой геопривязки растра, разрешением ячейки, номером колонки и номером ряда ячейки на растре. Точкой геопривязки растра может быть нижний

левый или верхний левый угол растра; это условие учитывается программным продуктом.

Расстояние между двумя ячейками растра есть функция местоположения их средних точек и разрешения ячеек. В растровой модели в ГИС определение длин вертикальных или горизонтальных линий проводится путем подсчета количества ячеек, по которым проходит линия, и умножением их на размер одной ячейки. Если линия ориентирована по диагоналям ячеек, то необходимо выполнить произведение количества ячеек на размер ячейки и на $\sqrt{2}$.

Площадь выбранных объектов на растре вычисляется как произведение количества ячеек на площадь отдельной ячейки.

5.2.3 Контрольные вопросы и задания для самостоятельной работы

- 1) Как определить координаты точки пересечения двух прямых?
- 2) Как определить координаты центроидов и центров?
- 3) Как определить площадь полигона?

Раздел 5.3

ФУНКЦИИ ВЫБОРА ДАННЫХ

При исследовании пространственного набора данных, прежде всего можно выбрать определенные пространственные объекты, чтобы временно ограничить область исследования, используя функции выбора данных (Retrieval functions). Такие выделения могут быть сделаны на пространственном основании или на основании атрибутивных данных, связанных с пространственными объектами. Средством выбора данных являются запросы пространственного выбора (Spatial selection queries). Запросы пространственного выбора можно комбинировать или выполнять в некой определенной последовательности для получения конечного результата.

5.3.1 Интерактивный пространственный выбор данных

Интерактивный пространственный выбор данных (Interactive spatial selection) выполняется путем указания на объекте или вычерчивания пространственных объектов на дисплее. После этого выбранные пространственные объекты индицируются на карте. Интерактивно определенные объекты называются селектированными или выбранными объектами. Затем ГИС выбирает пространственные объекты в активных слоях данных, которые накладываются на выбранные объекты.

Пространственные данные связаны с атрибутивными данными. Выбор пространственных объектов посредством этих связей приводит к выбору записей в таблицах. И наоборот, выбор записей в таблицах приводит к выбору пространственных объектов.

5.3.2 Пространственный выбор по атрибутивным условиям

Пространственные объекты могут быть выбраны путем формирования условий выбора на основе атрибутов пространственных объектов (Spatial selection by attribute conditions). Эти условия формулируются в форме запросов на языке структурированных запросов (Structured Query Language - SQL) (если данные находятся в реляционной базе данных) или на специфическом языке программного

обеспечения (если данные находятся в ГИС непосредственно). Выражения запросов могут быть простыми и комбинированными.

Для одного условия типа "где есть пространственные объекты с ...?" создается простое выражение, которое содержит последовательно название атрибута, оператор вычисления – арифметический (+, -, *, /) или сравнения (=, <, >, >=, <=, Like), значение атрибута. Например, условие "где есть земельные участки с площадью меньше 0,05 га." трансформируется в выражение запроса:

"площадь" < 0,05

Когда для выбора используется более одного критерия, создается составное выражение, которое содержит названия атрибутов, операторы вычислений, значения атрибутов и логические операторы (And, Or, Not). Например, запрос "где есть земельные участки с площадью менее 0,05 га. и тип использования земель жилая застройка" содержит два выражения с логическим оператором:

"площадь" < 0,05 And "тип" = 'жилая застройка'

5.3.3 Пространственный выбор на основании топологических отношений

Пространственный выбор на основании топологических отношений (Spatial selection using topological relationships) может выполняться определенным оператором. Выбор пространственных объектов зависит от их местоположения относительно других пространственных объектов. Наиболее общими являются следующие типы запросов:

- Выбор пространственных объектов, которые находятся внутри объектов (Selecting features that are inside selection objects). Этот тип запроса использует отношение включения (containment) между пространственными объектами. Очевидно, что полигоны могут включать полигоны, линии и точки, а линии могут включать линии и точки; другие включения невозможны.
- Выбор пространственных объектов, которые пересекают (Selecting features that intersect). Необходимый для выбора оператор будет выделять пространственный объект, который геометрически использует совместно общую часть объекта–источника.
- Выбор пространственных объектов, смежных с выбранными объектами (Selecting features adjacent to selection objects). Смежность есть отношение встречи, которое выражает то, что

пространственные объекты совместно используют границы. Оно применяется только для линейных или полигональных объектов.

- Выбор пространственных объектов по их удалению (Selecting features based on their distance). В качестве средства выбора пространственных объектов этот тип запроса использует функцию расстояния. Такой выбор может быть поиском внутри определенного расстояния от заданных объектов, или на заданном расстоянии, или больше заданного расстояния.

Эти запросы имеют разновидности операторов в списке в диалоговом окне "Выбор местоположением" (Select By Location).



Рис.5.3.1 – Пространственные отношения объектов

На рис.5.3.1 представлены пространственные отношения объекта В (светлого цвета) к объекту А (темного цвета):

- В "разъединен с" А (Disjoint);
- В "касается" А (Touches);
- В "накладывается на" А (Overlap);
- В "равен" А (Equal);
- В "покрывает" А (Cover);
- В "покрыт" А (Covered by);
- В "содержит" А (Contains);
- В "содержится в" А (Contains by).

5.3.4 Контрольные вопросы и задания для самостоятельной работы

- 1) Как выполняется пространственный выбор по атрибутивным условиям?
- 2) По каким топологическим отношениям формируются запросы пространственного выбора?

Раздел 5.4

ФУНКЦИИ КЛАССИФИКАЦИИ

5.4.1 Цели классификации по атрибутам

Эффективным инструментом пространственного анализа в ГИС являются функции автоматизированной классификации объектов по значениям их атрибутов (Classification functions). Функции классификации по атрибутам – это техника целенаправленного извлечения деталей из множества начальных данных с целью выявления закономерностей в пространственном распределении объектов и их визуализации.

Классификации по атрибутам могут быть:

- простыми, созданными на основе одного критерия (тип ландшафта);
- сложными, созданными на основе многих критериев одного покрытия (высота, количество осадков, экологические показатели и т.п.);
- комбинированными, созданными на основе многих параметров разных покрытий.

Набор начальных данных может быть результатом некоторой классификации; в таком случае говорят о повторной классификации (Reclassification). В процессе классификации входное множество значений атрибутов может быть оставлено неповрежденным.

В классификации векторных данных могут быть два возможных результата.

- 1) Входные объекты могут стать выходными объектами в новом слое данных с дополнительно назначенной категорией. Другими словами, пространственная протяженность оригинальных пространственных объектов не изменилась.
- 2) Выходные объекты получены в результате объединения соседних пространственных объектов с близкими значениями атрибутов. Такая постобрабатывающая функция называется пространственным слиянием, агрегированием или растворением.

5.4.2 Методы автоматизированной классификации по атрибутам

При автоматизированной классификации (Automatic classification) по атрибутам пользователь определяет атрибут классификации, метод классификации и число классов.

Выбор метода классификации опирается на оценку распределения данных посредством гистограммы. На горизонтальную ось гистограм выносятся значения, а на вертикальную – частота их появления в пределах данной выборки. Для автоматизированной классификации по атрибутам ArcMap предоставляет возможность использовать один из шести стандартных методов классификации [57].

Равный интервал

Метод равных интервалов (*Equal interval*) основан на следующем принципе классификации: каждый класс имеет равный диапазон значений. Разность между максимальным и минимальным значением одинакова для каждого класса. Программа вычитает минимальное значение в наборе данных из максимального значения. Полученное значение делит на заданное число классов. Затем получает предельное значение для первого класса путем прибавления результата деления к самому меньшему значению выборки. Таким же образом устанавливаются интервалы для остальной части классов. Равные интервалы более просты для понимания, так как диапазон для каждого класса одинаков.

Заданный интервал

Метод заданных интервалов (*Defined interval*) позволяет пользователю определить интервал, на который будет разделен весь диапазон значений атрибута. В отличие от схемы равных интервалов, где пользователь определяет количество интервалов, здесь необходимо указать значения интервалов. ГИС автоматически определяет количество классов на основании этого интервала.

Квантиль

Метод квантиль (*Quantile*) создает равное число объектов в каждом классе. ГИС упорядочивает объекты по принципу изменения их атрибута в интервале от максимального до минимального значения и суммирует их количество. Затем делит общее количество объектов

на число классов, которые задал пользователь. После этого присваивает первым по порядку объектам значения самого низкого класса, пока этот класс не будет заполнен, затем перемещается к следующему классу, заполняет его и так далее.

Естественная разбивка

Метод естественной разбивки (Natural breaks) основан на следующем принципе классификации: границы классов определяются там, где имеется резкий перепад между группами значений. По Jenks ГИС автоматически определяет максимальное и минимальное значение для каждого класса, используя математическую процедуру, которая анализирует резкие изменения в данных. Данная процедура выбирает интервалы, которые лучше всего группируют близкие значения, и максимизирует различия между классами.

Геометрический интервал

Метод геометрического интервала (Geometrical interval) основывается на интервалах классов, которые имеют геометрическую серию. Геометрическая серия - это последовательность значений, где каждое последующее значение получается умножением предыдущего значения на геометрический коэффициент. Алгоритм создает эти геометрические интервалы, минимизируя сумму квадратов элементов на класс. Это гарантирует, что диапазон каждого класса имеет приблизительно одно количество значений в каждом классе и что изменения между интервалами довольно последовательные.

Стандартное отклонение

Метод стандартного отклонения (Standard deviation) основан на следующем принципе классификации: каждый класс определен величиной отклонения от среднего по выборке. ГИС сначала находит среднее по выборке, разделив сумму всех значений на общее число объектов. После этого вычисляется среднеквадратическое отклонение путем вычитания среднего из каждого значения и возведения разности в квадрат. Полученные значения суммируются и делятся на число объектов. Из полученного выражения извлекается корень. Формула среднеквадратического отклонения каждого значения от среднего по выборке имеет вид:

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n}} \quad (5.4.1)$$

где S – среднеквадратичное отклонение, x – значение объекта, \bar{x} – среднее по выборке и n – число объектов.

5.4.3 Контрольные вопросы и задания для самостоятельной работы

- 1) Какие цели имеет классификация по атрибутам?
- 2) Приведите общую характеристику методов автоматизированной классификации по атрибутам.

Раздел 5.5

ОВЕРЛЕЙНЫЕ ФУНКЦИИ

5.5.1 Определение и общая характеристика оверлейных функций

Оверлей (Overlay) – это совместная обработка наложения двух или более исходных слоев одной географической области, в результате которой создается производный слой с новыми географическими данными как комбинация топологических сегментов исходных географических данных. Оверлей – это мощное средство анализа множества разноименных и разнотипных пространственных объектов.

Существует два основных пути выполнения оверлейных операций – на векторных моделях и на растровых моделях географических объектов. Геоинформационные системы предоставляют возможность использовать также комбинированный путь. Выбор метода зависит прежде всего от целей анализа, от того, какие данные уже существуют, от требуемой точности анализа, от сложности операций. Оверлейные операции могут приводить к отличающимся результатам. Каждый путь оверлейного анализа имеет свою специфику.

В системе, основанной на *векторных моделях*, топологические оверлейные операции являются более сложными, чем в системе, основанной на растровых моделях. Так как пространственные данные хранятся как точки, линии и/или полигоны, они требуют относительно сложных геометрических операций, чтобы вывести пересечение полигонов и создать новые узлы и дуги с объединенными значениями атрибутов.

Элементами оверлейных операций являются входной слой, оверлейный слой, выходной слой. Наложение пространственных объектов входных слоев позволяет разделять их на топологические сегменты и комбинировать из этих сегментов новые объекты в зависимости от цели анализа (Рис. 5.5.1).

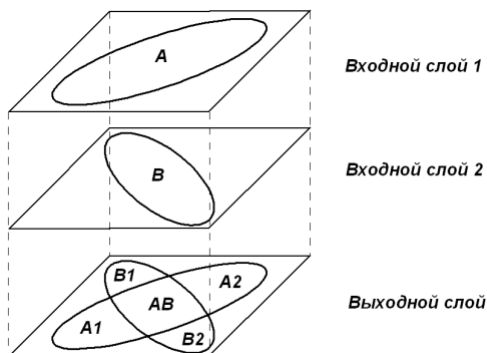


Рис. 5.5.1 - Концептуальное представление топологического наложения векторных слоев

Новые полигоны создаются на пересечении полигонов входного и оверлейного слоев. Наложение линейного объекта на полигональный объект разделяет его на два новых полигональных объекта. Новые объекты хранятся в выходном слое; входной слой не изменяется. Атрибуты пространственных объектов в оверлейном слое присваиваются соответствующим новым пространственным объектам вместе с атрибутами объектов входного слоя.

Во многих случаях требуются манипуляции с более чем двумя слоями векторных данных для достижения цели анализа. Операции выполняются ступенчатым способом: два входных слоя обрабатываются, чтобы сформировать производный слой; этот промежуточный слой затем обрабатывается с третьим слоем, чтобы сформировать следующий промежуточный слой, и так далее до достижения желаемого результирующего слоя карты.

В системе, основанной на *растровых моделях*, топологические оверлейные операции являются более простыми, чем в системе, основанной на векторных моделях. Каждая ячейка растрового слоя связана с одним соответствующим географическим местоположением. Это делает ее удобной для комбинирования характеристик многих слоев в одном слое. Обычно каждой характеристике присваиваются многие значения, позволяя пользователю математически комбинировать слои и назначать новые значения каждой ячейке в выходном слое.

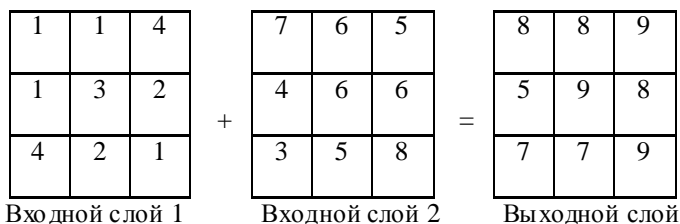


Рис. 5.5.2 – Иллюстрация суммирования данных при наложении растровых слоев.

5.5.2 Булева алгебра в топологическом наложении

Для реализации топологического наложения в ГИС используется алгебра логики. Это раздел математической логики, изучающий логические операции над высказываниями. Ее основоположником был английский математик Джордж Буль (G.BooI). В алгебре логики истинные значения высказываний принято обозначать числами 1 (истина – true) и 0 (ложь – false). Каждой логической операции соответствует функция, принимающая значения 1, 0. Такие функции называются функциями алгебры логики или булевыми функциями.

Чтобы определить, является ли определенное состояние истинным или ложным, в пространственном анализе используются логические операторы Булевой алгебры, которые обозначаются AND (И), OR (ИЛИ), NOT (НЕ) в текстовом формате, и соответственно \cap , \cup , \neg в символьном формате.

Два входных слоя топологического наложения можно рассматривать как два набора данных – набор А и набор В. Для них определяются следующие базовые логические операции:

- логическая операция *конъюнкция* $A \cap B$ – определяет *пересечение* двух наборов данных, идентифицирующее те сущности, которые принадлежат и набору А и набору В (истинно А и В);
- логическая операция *дизъюнкция* $A \cup B$ – определяет *объединение* двух наборов данных, идентифицирующее те сущности, которые принадлежат или набору А или набору В (истинно А или В);
- логическая операция *отрицание* $A \neg B$ – определяет *разность* двух наборов данных, идентифицирующая те объекты которые принадлежат А, но не В. (истинно не В).

Эти соотношения можно визуализировать с помощью диаграмм Венна.

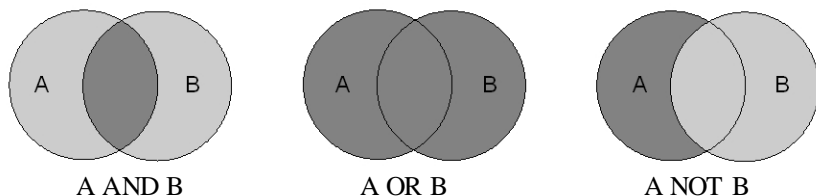


Рис. 5.5.3 – Диаграммы Венна

Используя базовые логические операции можно описать сложные логические функции.

Булева алгебра применяется в вычислении или моделировании новых объектов в топологической оверлейной обработке для систем, основанных на векторных и растровых моделях. Эти операции могут применяться ко всем типам данных – булевым, относительным, интервальным, порядковым или номинальным.

5.5.3 Векторные оверлейные операторы

5.5.3.1 Классификация векторных оверлейных операций

Топологические векторные оверлейные операции можно классифицировать по двум основаниям [58]:

- 1) по элементам, которые содержатся в слоях для наложения (или по слоям, содержащим точечные, линейные или полигональные элементы),
- 2) по типу операции (например, генерирование выходного слоя операциями соединения, пересечения или другими булевыми операциями).

Классификация векторных оверлейных операций по типам элементов каждого слоя.

Векторные оверлейные операции включают наложение точечных, линейных или полигональных пространственных объектов одного слоя на полигональные пространственные объекты другого слоя.

Следующая таблица 5.5.1 идентифицирует оверлейные опции для каждой возможной комбинации типов элементов, содержащихся в двух входных слоях.

Табл. 5.5.1 – Классификация векторных оверлейных операций по типам элементов

Тип элементов	Точки	Линии	Полигоны
Точки	Наложение точек (Point-on-Point)	Точка на линии (Point-in-Line)	Точка в полигоне (Point-in-Polygon)
Линии	Точка на линии (Point-in-Line)	Наложение линий (Line-on-Line)	Линия в полигоне (Line-in-Polygon)
Полигоны	Точка в полигоне (Point-in-Polygon)	Линия в полигоне (Line-in-Polygon)	Наложение полигонов (Polygon-on-Polygon)

Классификация векторных оверлейных операций по типу операции

Тип оверлейной операции определяется комбинацией использования логических операторов. При этом каждая комбинация будет производить различный выходной слой. Разные программные продукты ГИС могут иметь свои названия операций. Например, ArcGIS 9.x имеет инструменты оверлейных операций, приведенные в табл. 5.5.2.

Таблица 5.5.2 - Инструменты оверлейных операций в ArcGIS 9.x

Инструмент	Описание
Erase (Стирание)	Этот инструмент создает класс пространственных объектов из тех пространственных объектов или частей пространственных объектов, которые находятся вне стирающего класса пространственных объектов.
Identity (Идентичность)	Этот инструмент комбинирует части пространственных объектов, которые накладываются идентичными пространственными объектами для создания нового класса пространственных объектов.
Intersect (Пересечение)	Этот инструмент создает новый класс пространственных объектов из пересекающихся пространственных объектов общих в обоих классах пространственных объектов.
Spatial Join (Пространственное соединение)	Создает тип соединения таблиц, в котором поля из таблицы атрибутов одного слоя присоединены к таблице атрибутов другого слоя на основании относительного местоположения пространственных объектов в двух слоях.
Symmetrical Difference (Симметрическая разность)	Этот инструмент создает класс пространственных объектов из тех входных и оверлейных пространственных объектов или частей пространственных объектов, которые не накладываются.
Union (Объединение)	Этот инструмент создает новый класс пространственных объектов, комбинируя пространственные объекты и атрибуты каждого класса пространственных объектов.
Update (Обновление)	Этот инструмент обновляет атрибуты и геометрию входного класса пространственных объектов или слоя обновляющим классом пространственных объектов или слоем, которым они накладываются.

5.5.3.2 Алгоритмы векторных оверлейных операций

Из возможных комбинаций наложений точечных, линейных или полигональных пространственных объектов одного слоя на полигональные пространственные объекты другого слоя основными являются три следующих алгоритма обработки:

- "Точка в полигоне" (*Point-in-Polygon*),
- "Линия в полигоне" (*Line-in-Polygon*),
- "Полигон на полигоне" (*Polygon-on-Polygon*), иными словами "Оверлей полигонов" (*Polygon Overlay*).

Алгоритм операции "Точка в полигоне"

Операция "Точка в полигоне" (*Point-in-Polygon*) идентифицирует полигон, в который падает каждая точка. Результатом наложения "Точка в полигоне" является множество точек с дополнительными атрибутами полигона, в пределах которого находятся точки.

Алгоритм операции "Точка в полигоне" следующий:

- 1) Вначале используется экстенс полигона (минимальный ограничительный прямоугольник). Если точка лежит вне экстенса полигона, тогда она также должна находиться вне полигона. В таком случае анализ закончен. (Пример на Рис. а) Однако, если точка падает внутрь экстенса, требуется продолжение обработки.
- 2) В одном направлении от выбранной точки вычерчивается линия параллельно оси или X или Y, которая представляет собой направленный луч, называемый иногда "половиной линии".
- 3) Затем рассчитывается число пересечений этой "половины линии" с границей полигона. Если результат – четное число, то это указывает на то, что точка находится вне полигона. Если результат – нечетное число, то это указывает на то, что точка падает внутрь полигона.

Описанный алгоритм анализа "Точка в полигоне" работает также и для частных случаев: а) островных полигонов, б) полигонов с дырами, в) вогнутых полигонов. Примеры этих отдельных случаев показаны на рис.5.5.4.

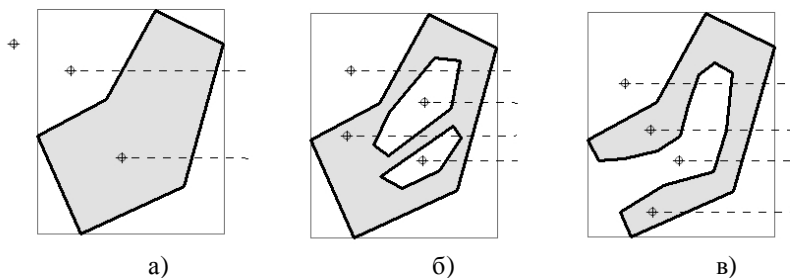


Рис. 5.5.4 – Наложение "Точка в полигоне"

Проблемы возникают, если а) точка находится точно на границе, б) точка находится на узле или вершине, в) сегмент линии коллинеарный "половине линии". В этих случаях, индивидуальные пакеты ГИС используют их собственные правила решения.

Пример использования анализа "Точка в полигоне": определение числа правонарушений, зафиксированных в точечном слое, по милицейским участкам.

Алгоритм операции "Линия в полигоне"

Полигональные пространственные объекты одного входного слоя могут быть наложены на линии (дуги) другого входного слоя. Линия может состоять из многих сегментов. Анализ "Линия в полигоне" идентифицирует полигон, который содержит каждую линию или сегменты линии. Результатом наложения "Линия в полигоне" является новый слой, содержащий линии с дополнительными атрибутами полигонов, в которые падают линии.

Поскольку линии и полигоны составлены из сегментов, анализ "Линия в полигоне" требует определения: пересекается ли любой из этих сегментов с накладываемой линией. Задача определения, пересекаются ли два сегмента линии, решается по ранее приведенным математическим зависимостям а разделе 5.2.

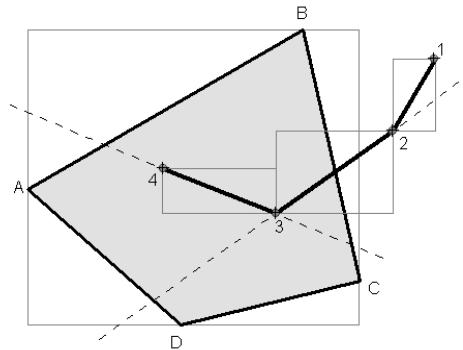


Рис.5.5.5 – Наложение "Линия в полигоне"

Геоинформационные системы используют следующий алгоритм.

- 1) Чтобы уменьшить число требуемых вычислений используются экстенды элементов. Если экстенд линии полностью находится вне пределов полигона, линия определенно находится вне пределов полигона, и на этом анализ заканчивается. В противном случае требуется продолжение обработки (например, рис. 5.5.5).
- 2) Поскольку линия может быть составлена из многих линейных сегментов, каждый сегмент линии должен быть испытан на пересечение или вложение в пределах полигона.
- 3) Полигоны могут иметь вогнутости или дыры внутри них, поэтому не достаточно определить, лежат ли оба конца сегмента линии в пределах полигона. Чтобы решить эту задачу, оба и полигон и сегмент линии *поворачивают* так, чтобы сегмент линии стал параллельным к одной оси.
- 4) Чтобы определить, находится ли каждый конец сегмента внутри или вне полигона, от каждого конца сегмента вычерчивается линия параллельно оси или X или Y, которая представляет собой направленный луч. Затем используется критерий "половины линии" (описанный в анализе "Точка в полигоне") относительно количества точек пересечения луча с границей полигона. Следует заметить, что точки пересечения "половины линии" не обязательно могут быть точками пересечения сегмента.
- 5) Если результат тестирования посредством "половины линии" показывает, что обе точки сегмента находятся внутри полигона и нет пересечений сегмента, то вся линия находится внутри полигона. Иначе, если начальная точка сегмента расположена вне полигона и первая часть сегмента линии находится вне полигона

до первой точки пересечения сегмента, вторая часть сегмента линии расположена внутри полигона до следующей точки пересечения сегмента, и так далее.

- б) Эта процедура применяется для каждого из сегментов линии, формирующих линию, которая анализируется.

Пример использования анализа "Линия в полигоне": определение локализации трубопроводов в пределах кварталов города.

Алгоритм операции "Полигон на полигоне"

Процесс "Полигон на полигоне" совместно обрабатывает накладываемые полигоны от двух входных слоев, чтобы создать новые полигоны в выходном слое. Результатом анализа "Полигон на полигоне" является выходной слой, содержащий новые полигоны с атрибутами от каждого из двух оверлейных полигонов.

Поскольку границы полигонов образованы из сегментов линий, анализ "Полигон на полигоне" требует определения того, пересекаются ли эти сегменты оверлейных линий. Обработка для анализа "Полигон на полигоне" является по существу той же самой, что анализ "Линия в полигоне".

5.5.3.3 Базовые оверлейные операций векторных моделей

Векторная оверлейная обработка использует предварительно упомянутые булевы логические операции AND, OR, и NOT, чтобы определить *пересечение* (Intersection) или *объединение* (Union) двух слоев, и *дополнение* (Complement) одного слоя в систему.

Векторная оверлейная обработка также использует булевы отношения *включения* (Inclusion (обозначение \subset)) и *исключения* (Exclusion (обозначение $\not\subset$)), чтобы определить, которая часть слоя или содержится или не содержится в пределах другого слоя в системе.

Например, рассмотрим два слоя A, B в ГИС:

- Операция " $A \subset B$ " определяет часть слоя A, которая *содержится* в слое B.
- Операция " $A \not\subset B$ " определяет часть слоя A, которая *не содержится* в слое B.

Существуют многочисленные векторные оверлейные операции, однако все они являются производными от четырех базовых операций: Union, Intersection, Inclusion, Exclusion [58].

Объединение

Объединение (Union) аналогично булеву логическому оператору OR (ИЛИ), где все элементы от обоих входных слоев будут присутствовать в выходном слое. Этот инструмент строит новый класс пространственных объектов комбинированием пространственных объектов и атрибутов каждого класса пространственных объектов. Объекты входного слоя разбиваются пересекающимися их объектами оверлейного слоя. Атрибуты объектов входного слоя содержат атрибуты объектов входного и оверлейного слоя.

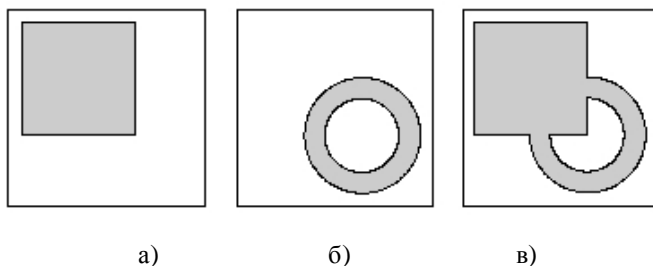


Рис. 5.5.6 – Объединение полигональных объектов двух слоев:
а) входной слой 1, б) оверлейный слой 2, в) выходной слой 3

Операция объединения двух слоев может быть представлена в булевой алгебраической форме:

$$\text{Слой 1} \cup \text{Слой 2} = \text{Слой 3} \quad (5.5.1)$$

Пересечение

Пересечение (Intersection) аналогично булеву логическому оператору AND (И), вычисляет геометрическое пересечение входных и оверлейных объектов. Объекты или части объектов, общие для всех слоев и/или классов объектов, будут записаны в результирующий класс. Объекты входной карты, которые не покрыты объектами оверлейной карты, игнорируются. Атрибуты полигонов выходной карты содержат атрибуты полигонов входной и оверлейной карт.

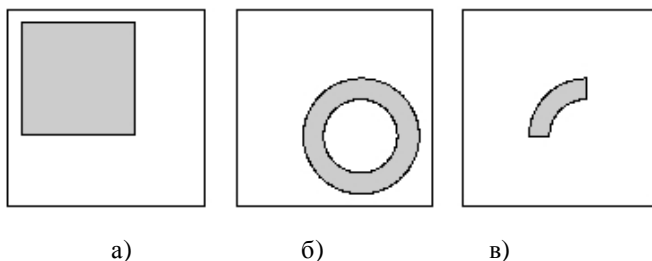


Рис.5.5.7 – Пересечение полигональных объектов двух слоев:
 а) входной слой 1, б) оверлейный слой 2, в) выходной слой 3

Операция пересечения двух слоев двух полигональных слоев может быть представлена в булевой алгебраической форме:

$$\text{Слой 1} \cap \text{Слой 2} = \text{Слой 3} \quad (5.5.2)$$

Включение

Операция *включения* (Inclusion) определяет часть оверлейного слоя, который содержится во входном слое. Выходной слой будет содержать все элементы первого входного слоя, а также все элементы второго входного слоя, существующие в географическом пространстве первого входного слоя. Включение использует входной слой как шаблон, в котором объекты оверлейного слоя отсекаются по ребрам и объединяются

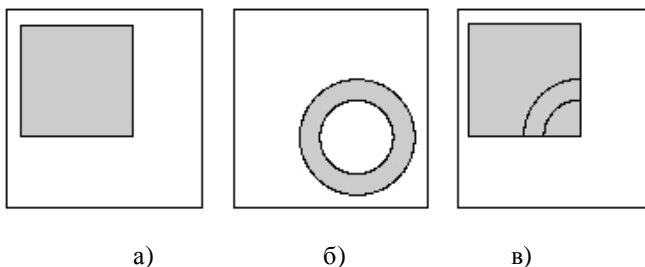


Рис. 5.5.8 – Включение полигональных объектов двух слоев
 а) входной слой 1, б) оверлейный слой 2, в) выходной слой 3

Операция включения двух слоев двух полигональных слоев может быть представлена в булевой алгебраической форме:

$$(\text{Слой } 2 \subset \text{Слой } 1) \cup \text{Слой } 1 = \text{Слой } 3 \quad (5.5.3)$$

Исключение

Операция *исключения* (Exclusion) “ $A \not\subset B$ ” определяет часть входного слоя, которая *не содержится* в оверлейном слое. Выходной слой будет содержать только те элементы первого входного слоя, которые не являются географическим пространством второго входного слоя. Исключение использует оверлейный слой как шаблон, которым объекты входного слоя отсекаются по его ребрам.

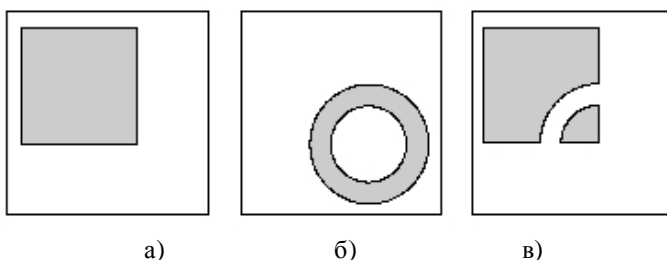


Рис. 5.5.9 – Исключение полигональных объектов двух слоев:
а) входной слой 1, б) оверлейный слой 2, в) выходной слой 3

Операция исключения двух слоев двух полигональных слоев может быть представлена в булевой алгебраической форме:

$$\text{Слой } 1 \not\subset \text{Слой } 2 = \text{Слой } 3 \quad (5.5.4)$$

5.5.4 Растровые оверлейные операторы

Векторные оверлейные операторы приводят к большей точности результата, но они геометрически сложные, и это иногда может быть следствием относительно медленной работы оператора. Оверлейные программы с растрами (Raster overlay operators) не имеют этого недостатка, так как большинство их выполняют вычисления "ячейки на ячейке", и поэтому они быстрые. Руководящим принципом является

сравнение или комбинирование значений характеристик местоположения, получаемых из двух слоев.

Обработка растров выполняется посредством языка представления операций на растрах. Такой язык используется в калькуляторе растров (Raster Calculator). Он позволяет вычислить новые растры на основании существующих, используя ряд функций и операторов. К сожалению, исчисления растра проходят скрытыми под различными названиями и не все предлагают ту же функциональность. Исчисления растров опирается на алгебру карт. При производстве нового растра нужно обеспечить имя для него и определить выражение вычисления. Это делается в назначении оператора следующего формата:

Output_raster_name := Raster_calculus_expression (5.5.5)

(Имя выходного растра := Выражение вычисления растра)

Выражение справа оценивается программой и растр, в котором затем хранятся результаты, называется именем с левой стороны.

Выражение может содержать ссылки на существующие растры, операторы и функции. Имена растра и константы, которые используются в выражении, называются его операндами. Когда выражение создано, программа будет вычислять наложения "ячейки на ячейке", начиная от первой ячейки в первом ряду, и продолжая до последней ячейки в последнем ряду.

Программы ГИС имеют широкий диапазон операторов и функций, которые могут использоваться в исчислении растров.

Арифметические операторы

Выражение вычисления растра может включать арифметические операторы (Arithmetic operators) +, -, *, /. Например, вычисление растра на рис. 5.5.10 выполняется на основании следующего выражения

$$C := A + B.$$

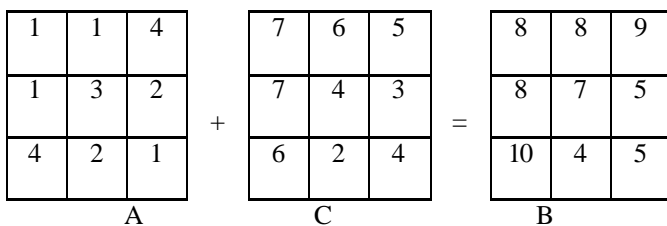


Рис. 5.5.10 Иллюстрация суммы двух растров

Выражение вычисления растра может также включать тригонометрические операторы *sin, cos, tan, asin, acos, atan*.

Выражение вычисления растра может включать несколько арифметических операторов.

Операторы сравнения и логические

Калькулятор растров может использовать операторы сравнения (Comparison operators) $<$, \leq , $=$, \geq , $>$, \diamond и логические операторы (Logical operators) AND, OR, NOT, XOR. Выражение $A \text{ XOR } B$ верно, если верно или A или B , но не оба. Пример применения оператора сравнения:

$$C := A \diamond B.$$

Операторы сравнения и логические вычисляют растры на истинностных значениях "истина" (true) и "ложь" (false).

Условные выражения

Условные выражения (Conditional expressions) в калькуляторе растров позволяют определить, где выполняется условие. Общий формат выражения:

$$\text{Output raster} := \text{IFF}(\text{condition}, \text{then expression}, \text{else expression}) \quad (5.5.6)$$

(Выходной растр := Условная функция (условие; выражение, где условие выполняется; выражение, где условие не выполняется))

Например:

$$\text{IFF}(4 = 5, \text{"land"}, \text{"lake"})$$

Комплексные выражения

Обработка растров может выполняться комплексными выражениями, включающими арифметические операторы, операторы сравнения и логические, на основании их табличных данных (Overlays using a decision table).

5.5.5 Контрольные вопросы и задания для самостоятельной работы

- 1) Приведите определение и общую характеристику оверлейных функций
- 2) Какие логические операции булевой алгебры могут быть для двух наборов данных?
- 3) Приведите классификации векторных оверлейных операций.
- 4) Опишите алгоритм операции "Точка в полигоне".
- 5) Опишите алгоритм операции "Линия в полигоне".
- 6) Приведите общую характеристику базовых оверлейных операций для векторных моделей.
- 7) Приведите общую характеристику базовых оверлейных операций для растровых моделей.

Раздел 5.6

ФУНКЦИИ ОКРЕСТНОСТИ

5.6.1 Определение окрестности

Под окрестностью (Neighborhood) географического объекта понимается окружающее его пространство. Во многих случаях пригодность некоторого местоположения для определенных целей зависит не только от того, что находится в нем, но и от того, что есть возле него. Поэтому, геоинформационные системы должны предоставить возможность оценить пространство, прилегающее к определенному месту, выполнить анализ близости (Proximity).

Функции окрестности (Neighborhood functions) определяют характеристики близости в окрестностях местоположения.

Чтобы выполнить анализ окрестности, необходимо:

- 1) установить, какое целевое местоположение представляет интерес и какая его пространственная протяженность;
- 2) решить, как определить окрестность для каждой цели;
- 3) определить, какие характеристики должны вычисляться для каждой окрестности.

Например,

- целью может быть торговый центр;
- его окрестность может быть определена как:
 - область в радиусе 2 км.,
 - область в пределах 10 минут езды по дорогам,
 - совокупность жилых зданий, для которых торговый центр самый близкий;
- характеристиками могут быть:
 - пространственное удаление,
 - количество людей, проживающих в окрестности торгового центра,
 - количество торговых объектов в окрестности.

В естественных средах возникают различные случаи расположений, окрестностей и особенностей соседства.

Современные программные продукты ГИС позволяют выполнить анализ окрестности, как на векторных моделях, так и на растровых моделях.

5.6.2 Операции окрестности в векторных моделях

Вычисления близости используют геометрическую дистанцию для определения окрестности одного или более географических объектов. Наиболее общим и полезным методом является генерирование буферных зон (*Buffer Zone Generation*). Другим методом, базирующимся на геометрической дистанции, является генерирование полигонов Тиссена (*Thiessen Polygon Generation*).

5.6.2.1 Генерирование буферных зон

Определение и назначение буферов

Буфер (*Buffer*) – это зона точно установленной ширины вокруг точечного, линейного или полигонального пространственного объекта. В результате генерирования буфера создается новый полигональный объект – буферная зона. Задача построения буферных зон требует определения геометрического места точек плоскости, удалённых от множества объектов не более чем на заданное расстояние.

Пространственный буфер может быть сгенерирован для анализа близости. Его используют для того, чтобы определить области и объекты, которые находятся или внутри или вне определенной буферной зоны.

Буферизация точечных пространственных объектов

Буферизация точечных пространственных объектов – самая простая форма буферизации, поскольку процесс включает создание круглого полигона относительно каждой точки, радиус которого равен ширине буфера b .

Имеются два метода назначения ширины буфера. Первый использует *фиксированную ширину* буфера для всех точек слоя. Второй - каждой точке назначается индивидуальная ширина буфера, основанная на атрибутах этого слоя, хранящихся в таблице (то есть *взвешенная ширина*).

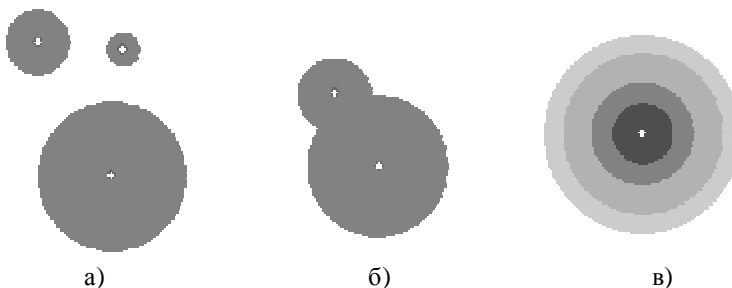


Рис. 5.6.1 - Буферизация точечных пространственных объектов: а) буферные зоны индивидуальной ширины, б) слитые буферные зоны, в) концентрические буферные зоны.

Если имеются множество точек в слое, подлежащем буферизации, то система должна проверить *перекрытия* буфера каждой точки. Любые наклаывающиеся части должны быть *удалены*. Поэтому результатом операции являются полигоны, представляющие области, покрытые всеми наклаывающимися буферами. Этот процесс включает две дополнительные операции: пересечение буферов и растворение общих границ буферов.

Буферизация линейных пространственных объектов

Алгоритм для буферизации линейных пространственных объектов более сложный, чем для буферизации точечных данных, так как линии могут быть составлены из множественных сегментов. Процесс для буферизации линейного пространственного объекта следующий.

1. Каждому сегменту полилинии назначают соответствующую ширину буфера (который может быть фиксированным для всех линии или взвешенным). Ширину буфера называют "буферным " расстоянием b .
2. Каждый сегмент полилинии имеет начальную вершину (X_1, Y_1) и конечную вершину (X_2, Y_2) . Используя эти координаты, вычисляют разности координат Δx и Δy между этими двумя конечными точками.
3. Определяются координаты конечных точек для параллельных буферных линий каждого сегмента линии с обеих сторон сегмента на перпендикуляре длиной b по формулам:

$$x_1 = X_1 \pm b * \cos [\text{arc tg} (\Delta y / \Delta x)] \quad (5.6.1)$$

$$y_1 = Y_1 \pm b * \sin [\text{arc tg} (\Delta y / \Delta x)] \quad (5.6.2)$$

и

$$x_2 = X_2 \pm b * \cos [\text{arc tg} (\Delta y / \Delta x)] \quad (5.6.3)$$

$$y_2 = Y_2 \pm b * \sin [\text{arc tg} (\Delta y / \Delta x)] \quad (5.6.4)$$

4. По полученным координатам определяются уравнения граничных линий буфера.
5. Рассчитываются пересечения граничных линий буфера для смежных сегментов, и назначаются координаты точек пересечения.
6. В каждой вершине линии граничные прямые линии буфера сопрягаются круговыми кривыми; в начальной и конечной точках полилинии граничные прямые линии сопрягают полукругом.
7. Отдельные части буфера полилинии сливаются в один буферный полигон

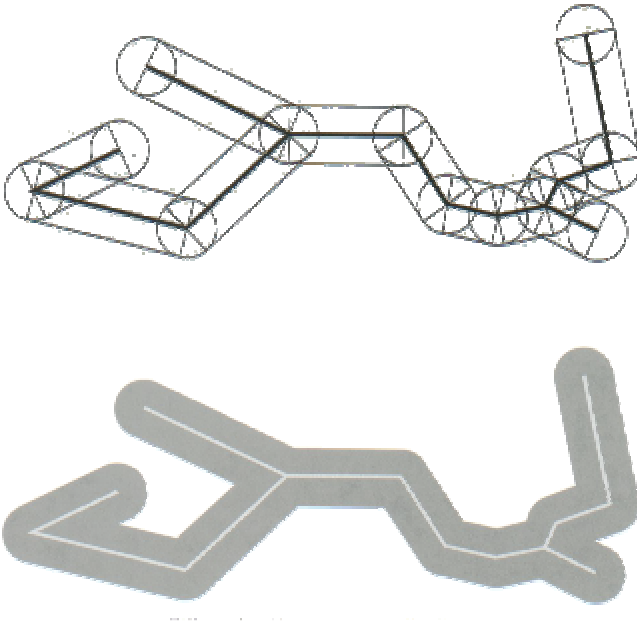


Рис. 5.6.2 – Буферизация линейных пространственных объектов [56].

Различные пакеты программного обеспечения ГИС могут определять концы буфера разными методами:

1. Сопряжение начальной точки и конечной точки линии с "полукруговыми" полигонами буфера радиуса b .
2. Простое отсечение параллельных буферных концов.
3. Сужение параллельных буферных концов, чтобы сблизить конечные точки.

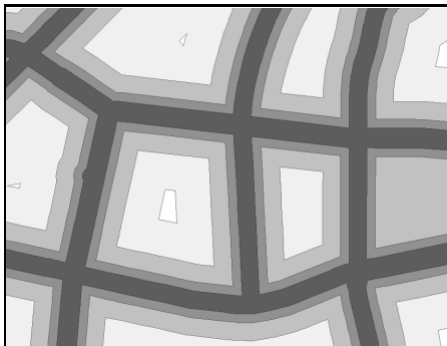


Рис. 5.6.3 – Зоны улично-дорожной сети

Если имеется множество линий в исходном слое, то система должна проверить перекрытия буферов для каждой линии. Любые накладывающиеся сечения должны быть удалены так, чтобы результатом операции были полигоны, представляющие область, покрытую всеми буферами. Этот процесс включает две дополнительные операции: пересечение буферов и растворение общих границ буферов.

Процесс буферизации приводит к новому слою в системе, состоящему из полигональных данных, которые представляют буферные зоны. Таблица результирующего полигона будет иметь идентификаторы для каждого полигона, созданного в процессе буферизации, и дополнительные атрибуты: ширина буфера и указания, находится ли полигон внутри буферной зоны или вне буферной зоны.

Буферизация полигональных пространственных объектов

Алгоритм буферизации полигональных пространственных объектов использует тот же самый процесс, как алгоритм буферизации

линий, с одним небольшим отличием – буферный полигон создан только на одной стороне линии, которая определяет полигон. По умолчанию метод состоит в том, чтобы создать буфер *вне* полигона, который окружает границу полигона. Некоторые пакеты программного обеспечения ГИС предоставляют также опции, чтобы создать буфер, который находится *внутри* границы полигона. На рис. 5.6.4 представлен пример возможных исходов буферизации полигона.

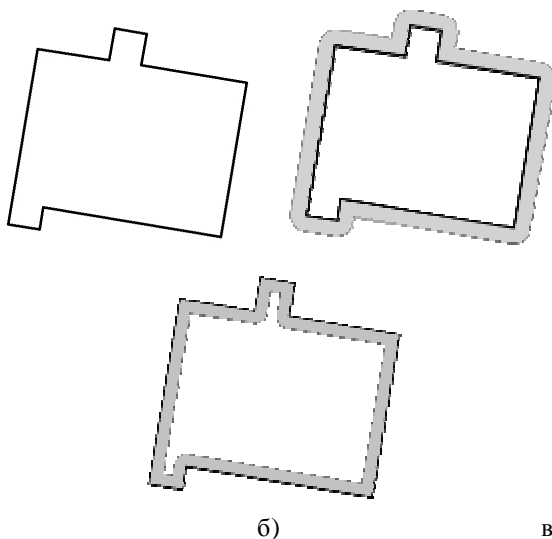


Рис.5.6.4 – Буферизация площадных пространственных объектов: а) исходный полигон, б) внешний буфер, в) внутренний буфер



Рис. - 5.6.5 Прибрежные защитные полосы водных объектов.

После построения буферов выполняются операции выбора объектов, находящихся в буферных зонах, и анализа их атрибутов.

5.6.2.2 Генерирование полигонов Тиссена

Другой метод использования геометрического расстояния для определения окрестности пространственно распределенных точек заключается в генерировании полигонов Тиссена (Thiessen Polygon Generation).

Полигоны Тиссена (Thiessen (1912)) – Вороного (Voronoi (1909)) – Дрихле (Dirichlet (1850)) - это полигональные области, образуемые на заданном множестве точек таким образом, что расстояние от любой точки области до данной точки меньше, чем для любой другой точки множества. Задача построения зон близости требует определения всех точек плоскости, для которых расстояние до объектов множества является минимальным.

Границами полигонов Тиссена являются отрезки перпендикуляров, восстановленных к серединам сторон треугольников в триангуляции Делоне, которая построена относительно того же точечного множества. Полигонам присваиваются атрибуты точечных объектов.

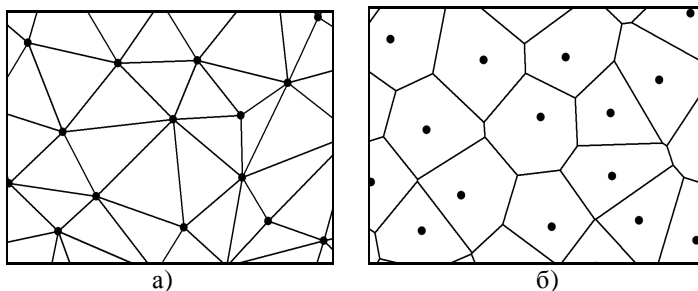


Рис. 5.6.6 – Генерирование полигонов Тиссена: а) триангуляция Делоне, б) полигоны Тиссена.

Полигоны Тиссена разделяют территорию на области близости к заданным точам, и поэтому могут рассматриваться в качестве окрестностей этих точек.

5.6.3 Операции окрестности в растровых моделях

Операции окрестности в растровых моделях становятся особенно важным средством, когда ситуация требует в большей мере анализа зависимостей между локализациями, чем интерпретации характеристик в индивидуальных локализациях.

5.6.3.1 Функции фокальной статистики

Функции Фокальные статистики (Focal Statistics) вычисляет статистические характеристики растровых данных в выделенной окрестности.

Сканированная ячейка и сканированная окрестность

В методе сканирования окрестности [58] операции окрестности на растровых моделях обычно называются "фокальными функциями", поскольку каждая выполненная операция генерирует значение для "фокуса" окрестности. Фокус окрестности вообще называют *сканированной ячейкой*, а ее окружающие соседние ячейки называют *сканированной окрестностью*. Сканированная окрестность может иметь различные размеры и формы, которые определяются выбором соответствующей опции в ГИС–пакете. Наиболее общие формы окрестности: квадрат, круг, кольцо, клин (рис. 5.6.7).

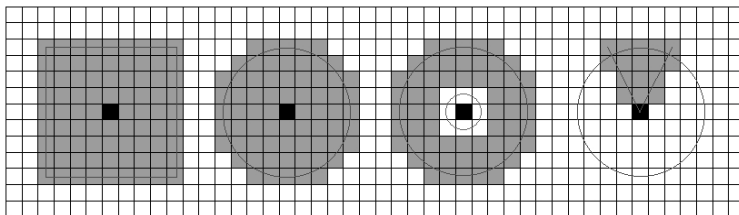


Рис.5.6.7 – Сканированные окрестности квадрат, круг, кольцо, клин .

В приведенных примерах на рис. 5.6.7 сканированная ячейка заштрихована темносерым цветом, а ячейки окрестности светлосерым цветом. Сканированная окрестность для каждой из этих форм окрестности включает сканированную ячейку, за исключением островной окрестности.

Пользователь ГИС может определять любую форму и размер для окрестности, в зависимости от задачи.

Процесс сканирования окрестности

Операции окрестности работают, перемещаясь по растровой карте по одной ячейке. Как только ячейка посещена, она становится сканированной ячейкой. Для этой ячейки вычисляется новое значение как функция ее сканированной окрестности. Все вычисленные значения затем помещаются в соответствующие ячейки выходной карты.

Статистический анализ окрестности

Для анализа могут использоваться различные статистические характеристики, чтобы характеризовать окрестность сканированной ячейки. Статистика, которая используется в операциях окрестности, будет зависеть от того, какие типы данных изучаются:

- данные отношений (*Ratio*),
- интервальные данные (*Interval*),
- порядковые данные (*Ordinal*),
- номинальные данные (*Nominal*).

Модуль статистического анализа окрестности выполняет девять операций.

- 1) *Сумма (Sum)*. Эта операция складывает значения сканированной ячейки и ее соседей (то есть, ячейки в ее сканированной окрестности), и хранит эту сумму в выходной теме. Данные входной карты могут быть отношением или интервалом.
- 2) *Среднее (Average)*. Эта статистическая операция вычисляет среднюю величину данных, существующих в сканированной окрестности. Данные входной карты могут быть отношением или интервалом.
- 3) *Максимум (Maximum)*. Эта операция присваивает значение ячейке сканированной окрестности, которое имеет наибольшее значение. Данные входной карты могут быть отношением, или интервалом, или порядковым номером.
- 4) *Минимум (Minimum)*. Эта операция присваивает значение ячейки в сканированной окрестности с самым низким значением. Данные входной карты могут быть отношением, или интервалом, или порядковым номером.

- 5) *Медиана (Median)*. Эта операция присваивает срединное значение сканированных значений окрестности. Данные входной карты могут быть отношением, или интервалом, или порядковым номером.
- 6) *Большая часть (Majority)*. Эта операция присваивает значение, которое повторяется наиболее часто в сканированной окрестности. Данные входной карты могут быть отношением или интервалом или порядковым номером или номиналом.
- 7) *Меньшая часть (Minority)*. Эта операция присваивает значение, которое повторяется наименее часто в сканированной окрестности. Данные входной карты могут быть отношением, или интервалом, или порядковым номером, или номиналом.
- 8) *Разнообразие (Diversity)*. Эта операция присваивает значение, равное числу различных значений в сканированной окрестности. Данные входной карты могут быть отношением, или интервалом, или порядковым номером, или номиналом.
- 9) *Амплитуда (Range)*. Эта операция присваивает значение, равное числу различных значений в сканированной окрестности. Данные входной карты могут быть отношением или интервалом или порядковым номером или номиналом.

Пример сканированной окрестности представлен на рис.5.6.8.

3	7	16
7	16	11
11	16	25

Рис. 5.6.8 – Пример сканированной окрестности

Для этого примера статистические операции над сканированной ячейкой и ее окрестностью представлены результатами:

Сумма	= 112	
Среднее	= 12,44	(112 / 9)
Максимум	= 25	
Минимум	= 3	
Медиана	= 11	(3,7,7,11,11,16,16,16,25)
Большая часть	= 16	
Меньшая часть	= 25	
Разнообразие	= 5	(3,7,11,16,25)
Амплитуда	= 22	(25 – 3)

5.6.3.2 Функции распространения

Определение окрестности одного или большего количества целевых мест (источников) может зависеть не только от расстояния, но и также от *направления и отличий* местности в различных направлениях. Обычно целевое местоположение содержит "исходный материал", который распространяется через какое-то время. Этот "исходный материал" может быть воздухом, водой или загрязнением земли, выходящими пассажирами поезда, весенними водами, радиоволнами от ретрансляционной станции.

Продвинутые методы анализа окрестности заключаются в вычислении распространения (Spread computations). Функции распространения основаны на предположении, что явление распространяется во всех направлениях, хотя не обязательно одинаково во всех направлениях. Поэтому, они используют локальные особенности местности, чтобы вычислить локальное сопротивление распространению. Локальные факторы местности влияют на распространение, делая его легким или более сложным. Многие ГИС-программы обеспечивают поддержку этого вида вычисления распространения.

В ArcGIS 9.x этот вид анализа окрестности обеспечивают функции затратного расстояния (Cost distance functions). Здесь под термином "затраты" (Cost) в геопространственном анализе понимается любой вид прямых или относительных издержек на перемещение, выраженных в выбранных единицах – стоимости, длины, времени, энергии, сопротивления, импеданса и тому подобное.

Растр затрат

Растр затрат (Cost raster) определяет затраты на перемещения через каждую ячейку.

Вычисление распространения включает одно или несколько целевых мест, которые являются расположениями источника любого распространения. Источник представляет расположение объекта интереса.

В общем случае растр затрат создается на основании нескольких критериев. Например, для создания растра затрат территории для строительства дороги исходными данными могут быть: А – растр типов грунтов, значения ячеек которого являются кодами типов

грунтов, В – растр уклонов рельефа, значения ячеек которого выражены в процентах крутизны ската.

Исходные наборы данных могут быть в разных системах измерения. Поэтому для совместного использования их необходимо привести к общей шкале. Один из методов приведения к общей шкале есть переклассификация (Reclassifying) наборов данных. Например, для строительства дороги растр использования земель и растр уклонов рельефа могут быть переклассифицированы по шкале 1–10. При этом метод переклассификации для каждого растра определяется влиянием значения атрибута на относительные затраты в строительстве дороги.

Следующий шаг в создании растра затрат – есть объединение приведенных к общей шкале растров. При равном влиянии исходных наборов данных на результат выполняют простое сложение двух растров. При разном влиянии исходных наборов данных на результат выполняют взвешенное сложение двух растров. Например, для строительства дороги итоговый растр затрат может быть получен выражением:

$$C := A * pA + B * pB \quad (5.6.5)$$

где P_A, P_B – веса растров.

Функция Затратное расстояние

Функция *Затратное расстояние* (Cost distance) создает выходной растр расстояний в единицах накопленных затрат, в котором каждой ячейке присваивается значение, представляющее наименьшие накопленные затраты на перемещение от этой ячейки до ближайшего источника. При этом окрестность задается максимальным расстоянием от источника. Для вычисления выходного растра расстояний в единицах накопленных затрат требуются растр с источником и растр затрат.

Алгоритм вычисления использует представление ячейки узел/связь. В представлении узел/связь каждый центр ячейки считается узел, а каждый узел соединен с соседними узлами связями.

При переходе от ячейки к одной из четырех соседней ячейки затраты на перемещения по связям с соседними узлами получаются по формуле:

$$a1 = (cost1 + cost2) / 2 \quad (5.6.6)$$

где $cost1$ есть затраты ячейки 1, $cost2$ есть затраты ячейки 2 и $a1$ есть общие затраты связи из ячейки 1 в ячейку 2.

Если перемещение происходит по диагонали путь увеличивается в $1.414214 = \sqrt{2}$ раза, поэтому затраты перемещения по связи получаются по формуле:

$$a1 = 1,414214 (cost1 + cost2) / 2 \quad (5.6.7)$$

Алгоритм использует итеративный процесс, который начинается с исходных ячеек. В первой итерации ячейкам источника назначается ноль, поскольку нет накопленных затрат на возвращение к себе. Далее, все соседи ячеек–источников активизируются, и затраты назначаются связям между узлами ячеек–источников и узлами соседних ячеек с использованием указанных выше формул. В последующей итерации список активных соседних ячеек расширяется с использованием ячейки самых низких затрат.

Накопленные затраты (Accumulative cost) вдоль пути наименьших затрат вычисляются как простая сумма затрат на перемещения по связям от попарных ячеек окрестности:

$$Accum_cost = a1 + a2 + \dots \quad (5.6.8)$$

Полученное значение присваивается ячейке, соответствующей началу перемещения до ближайшего источника. Результатом этого процесса является растр накопленных затрат или взвешенного расстояния.

1	1	1	2	8	14.50	14.95	15.95	17.45	22.45
4	4	5	4	9	12.00	12.45	14.61	16.66	21.44
4	3	3	2	10	8.00	8.95	11.95	13.66	19.66
4	5	6	8	8	4.00	6.36	8.00	10.00	11.00
4	2	1	1	1	0.00	3.00	4.50	5.50	6.50

Рис. 5.6.9 – Вычисление распространения [54]

Функция Затратное расстояние находит широкое применение при определении пути наименьших затрат во многих приложениях, например, при проектировании трассы дороги или магистрального трубопровода.

Функция Направления наименьших затрат

В поисковых вычислениях (Seek computations) предполагается, что для возврата к источнику феномен выберет путь наименьших накопленных затрат, а не распространение во всех направлениях. Примером могут быть потоки дождевых осадков. Это требует вычисления предпочитаемого локального направления распространения.

Функция Направления наименьших затрат (Cost Back Link) в ArcGIS 9.x для каждой ячейки растра накопленных затрат вычисляет смежную ячейку с наименьшими накопленными затратами. Каждой ячейке присваивает код направления на смежную ячейку с наименьшими накопленными затратами в соответствии с системой кодирования на рис. 5.6.10

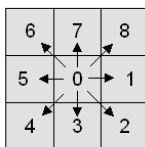


Рис.5.6.10 – Схема кодирования направлений

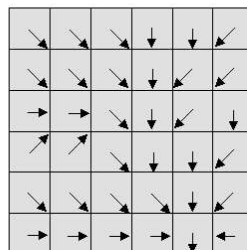
Иллюстрация функции направления наименьших затрат представлена на рис.5.6.11.

88	82	79	81	68	59
84	77	66	59	56	60
79	63	54	47	48	58
74	68	65	32	41	34
78	71	57	31	26	29
84	63	44	22	21	22

а)

2	2	2	3	3	4
2	2	2	3	4	4
1	1	2	3	4	3
8	8	2	3	3	4
2	2	2	2	3	4
1	1	1	1	3	5

б)



в)

Рис. 5.6.11 – Иллюстрация функции направления наименьших затрат:
а) растр накопленных затрат, б) растр кодов направлений, в) растр направлений

5.6.3.3 Функции распределения

Функция Распределение

Функция Распределение (Allocation function) создает выходной растр, в котором каждой ячейке присвоено значение ближайшей к ней ячейки-источника. Функция Распределение позволяет определить, какие ячейки, к какому источнику относятся, на основании значения по прямой.

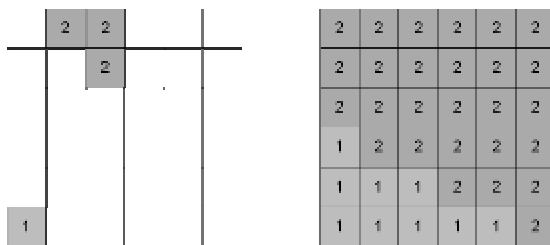


Рис.5.6.12 – Иллюстрация функции распределения: а) растр источник, б) растр распределения.

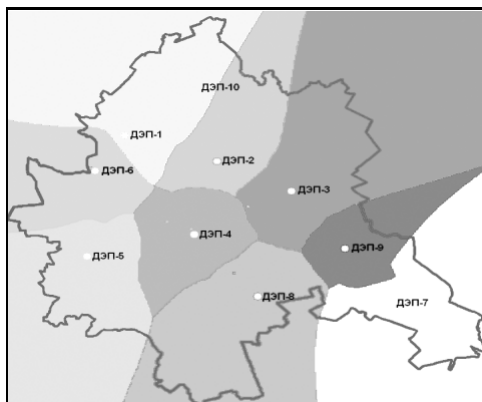


Рис. 5.6.13 – Пример распределения зон обслуживания улично-дорожной сети

Функция Распределение используется для определения ближайших объектов к объекту-источнику. Например, а) ближайших пожарных гидрантов к объекту пожаротушения, б) зон обслуживания

улично–дорожной сети к техническим базам дорожно–эксплуатационных управлений города (рис .5.6.13).

Функция Затратное распределение

Функция Затратное распределение (**Cost Allocation**) вычисляет для каждой ячейки ее ближайший источник по наименьшим накопленным затратам раstra накопленных затрат.

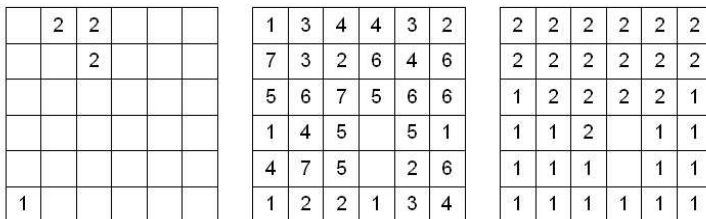


Рис. 5.6.14 – Иллюстрация функции затратного распределения:
 а) растр источник, б) растр затрат, в) растр распределения.

5.6.4 Контрольные вопросы и задания для самостоятельной работы

- 1) Что представляет собой окрестность?
- 2) Для чего выполняется буферизация точечных пространственных объектов?
- 3) Опишите алгоритм буферизации линейных пространственных объектов.
- 4) Опишите буферизацию полигональных пространственных объектов.
- 5) Для чего генерируются полигоны Тиссена?
- 6) Как выполняется статистический анализ окрестности?
- 7) Как создается растр затрат и растр затратных расстояний?
- 8) Как построить растр направлений наименьших затрат?
- 9) В чем состоит различие между функцией распределения и функцией затратного распределения?

Раздел 5.7

ФУНКЦИИ СВЯЗНОСТИ

Функции связности позволяют выполнять сетевой анализ.

5.7.1 Определение и характеристика сети

В геоинформационных системах *сеть* (Network) – это система связанных точками пространственных линейных объектов.

Сети представляют географические объекты для транспортирования вещества, энергии, информации, например, сеть дорог – автомобильных, железных; улично–дорожная сеть города; сети маршрутного транспорта, в том числе пассажирского; сети инженерной инфраструктуры города – водоснабжения, водоотведения, теплоснабжения, газоснабжения электроснабжения; гидрографическая сеть водотоков и так далее. По сетям может транспортироваться почти все: люди, машины, вещи, продукция, загрязнение воды в реках, телефонные сообщения в телефонной сети ...

На сетях решаются множество различных транспортных задач на основе сетевого анализа (Network analysis):

- 1) управление инфраструктурой и ее развитием на основе базы данных на все объекты транспортного процесса (анализ транспортных потоков и пассажиропотоков, планирование и анализ маршрутной сети, составление и анализ отчетов по ДТП), реструктурирование маршрутов, поддержка эксплуатации систем энергоснабжения, сигнализации и связи);
- 2) управление парком подвижных средств и логистика (нахождения оптимального маршрута, задача коммивояжера, транспортная задача - полномасштабная организация перевозок различных грузов из многих источников по многим адресам, мультимодальная транспортировка)
- 3) управление движением (слежение за транспортными средствами с помощью GPS диспетчеризация, увязка расписаний с другими видами транспорта)

Сетевой анализ может выполняться на векторных или растровых данных. Сети на векторных данных имеют следующие особенности.

- В векторных данных линейные пространственные объекты *естественно ассоциируются* с элементами географических

объектов и позволяют обрабатывать типичные транспортные характеристики, такие как пропускная способность и затраты на единицу. Кроме того, сетевой анализ на векторных данных позволяет получить более точные результаты.

- Сеть определяется топологически связанными ее элементами. Каждый линейный пространственный объект (дуга) имеет начальную и конечную точки. Дуги присоединяются одна к другой в узлах, формируя *связность* (Connectivity) сети. Связность сети делает возможным выполнение сетевого анализа. Связность сети используется также для определения меры сложности сети.
- Важнейшей особенностью любой сети является ее *направленность*. По направленности сети разделяют на транспортные и инженерные. Транспортные сети - это ненаправленные сети. Это означает, что движение по линии принципиально может быть в прямом и обратном направлении, хотя организация движения может предусматривать одностороннее движение. Транспортные сети моделируют используя сетевые наборы данных. Инженерные сети – это направленные сети. Потоки вещества перемещаются в одном направлении линии. Инженерные сети моделируют, используя геометрические сети.
- Сети могут быть *одноуровневыми* или *многоуровневыми*. Для многих приложений сетевого анализа используются одноуровневые (плановые) сети, рассматриваемые в двумерном пространстве, например, водные потоки. Плановые сети легче в обработке, так как они имеют топологические правила. Многоуровневые сети не рассматриваются как плановые, так как они имеют многоуровневые пересечения, например, тоннели, путепроводы, подземные переходы.

Математически сети описываются теорией графов, а решение сетевых задач выполняется средствами линейного программирования. Программными пакетами ГИС поддерживаются различные функции пространственного анализа на сетях. Базовыми классическими функциями являются:

- нахождение лучшего пути, которая генерирует путь наименьших затрат на сети между парой определенных мест на основании геометрических и атрибутивных данных;
- разделение сети, которая назначает элементам сети (узлам или сегментам) различные местоположения, используя предопределенные критерии.

Продвинутый сетевой анализ использует специальную модель данных - геометрическую сеть. Геометрическая сеть – это коллекция связанных ребер (Edges) и соединений (Junction) с правилами связности, которые используются для представления и моделирования поведения общей сетевой инфраструктуры в реальном мире. Ребра и соединения могут быть простыми и сложными. Примерами ребер являются оси дорог, трубопроводы, средние линии рек. Примерами сложных соединений являются перекрестки, переключатели, трансформаторы, распределительные станции.

Геометрические сети предоставляют расширенные возможности сетевого анализа, моделирования географических объектов для транспортирования, обеспечивают более совершенный уровень управления инженерной инфраструктурой.

5.7.2 Нахождение лучшего пути

Нахождение лучшего маршрута (Finding the best route) или нахождение оптимального пути (Optimal path finding) используется, когда требуется найти путь наименьших затрат между двумя узлами в сети. Решение задачи обеспечивает алгоритм Дейкстра (E. Dijkstra).

Лучший маршрут может быть наискорейшим, кратчайшим или определенным в соответствии с заданными затратами. Цель заключается в нахождении последовательности связанных линий от узла-источника до узла-назначения.

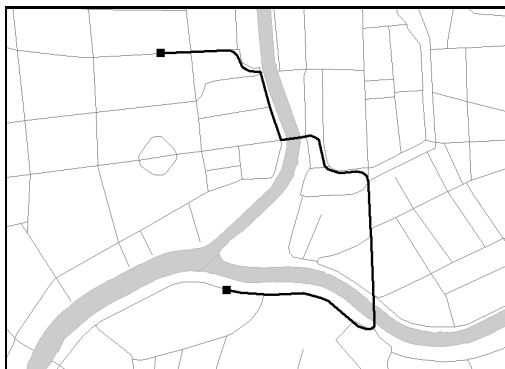


Рис. 5.7.1 – Кратчайший маршрут

Проблемы, связанные с поиском оптимального пути, могут иметь добавочное требование: *дополнительные узлы*, которые должны посещаться в пути. Это требование может иметь упорядоченный и неупорядоченный характер. В упорядоченном поиске оптимального пути определена последовательность, в которой эти добавочные узлы должны посещаться; в неупорядоченном поиске оптимального пути не определена последовательность посещения дополнительных пунктов.

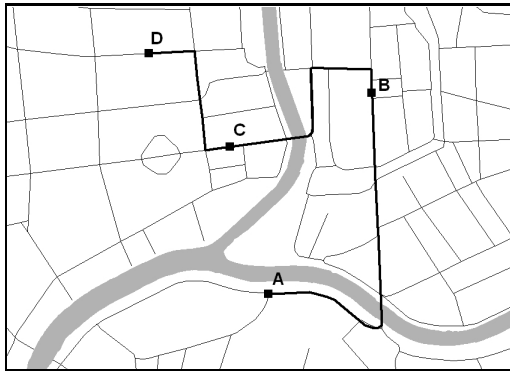


Рис. 5.7.2 - Маршрут А-Д с дополнительными узлами В,С.

Функция затрат может быть простой: например, она может определяться как общая длина всех линий на пути. Функция затрат может быть более комплексной, например, аргументами которой будут не только длины линий, но и пропускная способность, максимальный тариф перевозки и другие особенности. Возможен также детальный учет затрат на повороты в каждом узле - при входе на узел по одной линии и выходе с узла по другой, а также запрещенных (разрешенных) направлений движения.

5.7.3 Разделение сети

Разделение сети (Network partitioning) имеет целью назначение линиям и/или узлам сети некоторого количества целевых местоположений взаимноисключающим способом. К этой группе относятся функции *распределения сети* и *трассирования*.

Распределение сети

Обычно целевые расположения играют роль центра обслуживания для сети. Это может быть любой вид сервиса: поликлиническое обслуживание, школьное образование, снабжение водой. Тип разделения сети относительно центров обслуживания известен как Распределение сети.

Пусть имеется ряд целевых местоположений, которые функционируют как ресурсные центры. При *распределении сети* (Network allocation) проблема заключается в том, чтобы разделить сеть на части, которым назначить определенный центр обслуживания исключительно. В простой проблеме распределения сервисный центр назначается тем линиям или сегментам, по которым он самый близкий или находится в пределах заданного расстояния. В сложной проблеме распределения возникает необходимость учета дополнительных факторов, например, а) мощность центра (количество посещений поликлиники, количество школьников, киловатты), б) потребление ресурсов, которое может изменяться между линиями или сегментами линии. Следует также иметь в виду, что некоторые улицы отличаются количеством происшествий, проживающих там детей, объектов промышленности с высоким потреблением электричества, смежных земельных участков (Рис. 5.7.3).

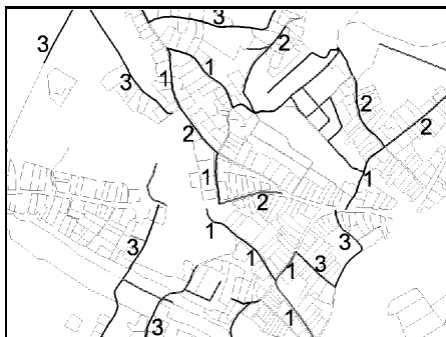


Рис. 5.7.3 - Категории улиц по количеству смежных земельных участков

Сеть обслуживания (Service net) любого центра - это подмножество дистрибьюторской сети, фактически связанная часть сети. Существуют различные методы, чтобы отнести сеть линий или

их сегменты к определенному центру. В ArcGIS 9.x эта задача решается функцией Область обслуживания (Service area) Рис. 5.7.4.

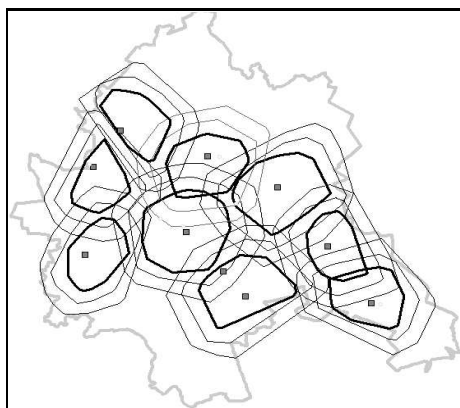


Рис. 5.7.4 - Области обслуживания, удаленные от дорожно-эксплуатационных управлений города по реальным путям на 3, 4, 5 км.



Рис. 5.7.5 - Сеть обслуживания дорог, удаленных по реальным путям до 3,5 км. от одного дорожно-эксплуатационного управления.

Трассирование

Сетевой анализ включает Трассирование (Tracing). Функция Трассирование по геометрической сети создает связанные наборы элементов сети, в соответствии с некоторым условием.

При помощи инструментов трассирования можно найти:

- все элементы сети, которые лежат вверх или вниз по течению от заданной точки на сети (рис. 5.6.6, 5.6.7);
- общие затраты всех элементов сети, которые лежат вверх по течению от заданной точки на сети;
- общие пространственные объекты, которые выше по течению от заданной точки;
- все объекты, которые связаны или не связаны с данной точкой через сеть;
- петли, которые могут быть результатом маршрутов между точками на сети.

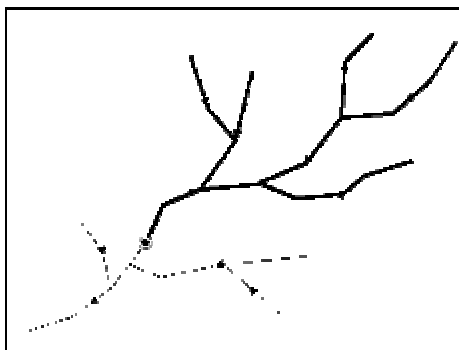


Рис. 5.7.6 – Трассирование по геометрической сети вверх по течению

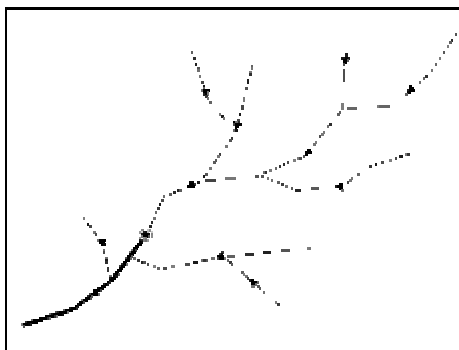


Рис. 5.7.7 – Трассирование по геометрической сети вниз по течению.

Трассирование выполняется, когда нужно узнать, какая часть сети условно соединяется с выбранным узлом на сети, известным как начало трассы. Для узла или линии быть условно соединенным

означает, что существует путь от узла/линии к началу трассы, и что путь соединения выполняет ряд условий. Какими являются эти условия, это зависит от приложения. Они могут включать направление пути, вместимость, длину, потребление ресурсов вдоль пути, и так далее. Условие обычно - это логическое выражение, например:

- путь должен направляться от узла/линии к началу трассы;
- его вместимость (определяемая как минимальная вместимость линий, которые составляют путь) должна быть ниже установленного порога;
- длина пути не должна превышать установленной максимальной длины.

Трассирование может быть полезной функцией для многих проблем, связанных с инженерными сетями. Например, определение вентилях вверх по течению при авариях в распределительных инженерных сетях.

5.7.4 Контрольные вопросы и задания для самостоятельной работы

- 10) Что представляет собой окрестность?
- 11) Для чего выполняется буферизация точечных пространственных объектов?
- 12) Опишите алгоритм буферизации линейных пространственных объектов.
- 13) Опишите буферизацию полигональных пространственных объектов.
- 14) Для чего генерируются полигоны Тиссена?
- 15) Как выполняется статистический анализ окрестности?
- 16) Как создается растр затрат и растр затратных расстояний?
- 17) Как построить растр направлений наименьших затрат?
- 18) В чем состоит различие между функцией распределения и функцией затратного распределения?

ПОСЛЕСЛОВИЕ

Материал учебного пособия представляет собой каркас теоретической части базового курса "Основы ГИС". В учебном пособии изложены устоявшиеся концепции, подходы, принципы построения и функционирования географических информационных систем, которые реализованы во многих развитых программных продуктах ГИС.

Изучение этого теоретического материала подлежит проработке и закреплению при практическом освоении программных пакетов ГИС и создании на их основе ГИС-проектов. Компания ESRI уделяет повышенное внимание популяризации ГИС-технологии и внедрению географического мышления в учебный процесс. По оценкам аналитиков известной компании Даратек, программное обеспечение ESRI занимает ведущие позиции в сфере образования. В соответствии со стратегией ESRI дистрибьюторы этой компании реализуют долгосрочную программу поддержки учебных заведений, направленную на развитие ГИС образования.

Дальнейшее наращивание знаний в области геоинформационных систем в рамках подготовки специалистов будет происходить путем изучения последующих профилирующих дисциплин учебного плана, о которых говорилось во введении.

Учебное пособие ориентировано на формирование у специалистов начальных знаний в области геоинформационных систем и технологий. Вместе с тем материал пособия может быть полезным как студентам других специальностей, так и специалистам, которые внедряют передовые информационные технологии в бизнес процессы, при решении многих управленческих задач. Этот подход наиболее плодотворен, когда он начинает прививаться с молодых лет и органично входит в сознание людей.

Сегодня эта технология является одним из наиболее популярных и полезных инструментов в научных исследованиях. ГИС - это универсальный инструмент исследователя. Функции пространственного анализа применяется в более чем 100 дисциплинах, охватывающих большинство направлений научных и прикладных исследований. ГИС также является прекрасным средством презентации результатов проведенных исследований.

В заключение следует обратить особое внимание на еще одну грань геоинформационных систем и технологий. ГИС - это не просто одна из современных информационных технологий. ГИС помогает сформировать у людей новый взгляд на мир, обеспечивающий его комплексное восприятие и лучшее понимание взаимосвязей между его составляющими. Это прогрессивный образ мышления, способ познания окружающего нас мира, инструмент, помогающий перестройке нашего мировоззрения.

БИБЛИОГРАФИЧЕСКОЕ ОПИСАНИЕ

Использованная литература

1. Саати Т. Аналитическое планирование. Организация систем / Т. Саати, К. Кернс ; пер. с англ. - М.: Радио и связь, 1991. - 224 с.
2. Берталанфи Людвиг Фон. Общая теория систем: обзор проблем и результатов/ Берталанфи Людвиг Фон // Системные исследования. - М.: Наука, 1969. – С. 30-54.
3. Философский энциклопедический словарь / Гл. редакция: Л.Ф.Ильичев, П.Н.Федосеев, С.М.Ковалев, В.Г.Панов. - М: Сов. энциклопедия, 1983. - 840 с.
4. Аверьянов А.Н. Системное познание мира: Методологические проблемы / А.Н. Аверьянов. - М.: Политиздат, 1985.- 263 с.
5. Проблемы общей теории систем как методологии / В.Н. Садовский // Системные исследования, 1973, С. 127-135
6. Оптнер, Станфорд Л. Системный анализ для решения деловых и промышленных проблем / Оптнер, Станфорд Л. - М.: Сов.радио, 1969. – 312 с.
7. Диалектика и системный анализ / Под. ред. Д.М.Гвишиани. - М.: Наука, 1986. - 336 с.
8. Про затвердження "Програми розвитку земельної реформи та управління земельними ресурсами в місті Харкові до 2006 року" / Рішення XVIII сесії IV созиву від 24.12.2003 р. Харківської міської ради Харківської області [Electronic Resource] .- URL: www.city.kharkov.ua
9. Гради Буч. Объектно-ориентированный анализ и проектирование с примерами приложений на С++. Второе издание / Гради Буч; пер. с англ; под ред. И. Романовского и Ф. Андреева.- Калифорния.: Rational Санта-Клара, 2006. - 380 с.
10. Клинтон У.Д., Президент США. Координация в области получения доступа к данным: национальная инфраструктура пространственных данных // Правительственное распоряжение, Белый дом, 11 апреля 1994 г.
11. ДеМерс М.Н. Географические информационные системы. Основы / ДеМерс М.Н.; пер. с англ. - М.: Дата+, 1999.- 491 с.
12. Goodchild, M.F. Geographical information science //International Journal of Geographical Information Systems, 1992, 6(1), p. 31-45.

13. The NCGIA Core Curriculum in GIScience / Goodchild, M.F., Kemp K.K., eds. - NCGIA University of California, Santa Barbara CA., 2000 [Electronic Resource] .- URL: <http://www.ncgia.ucsb.edu/>
14. Чоговадзе Г.Г. Информация: информация, общество, человек / Г.Г.Чоговадзе.- М.:ООО Дата+, 2003. -320 с., с.121-134
15. Кошкарев А.В. Геоинформатика / А.В.Кошкарев, В.С.Тикунов; под ред. Д.В.Лисицкого.- М.: "Картоцентр - Геодезиздат", 1993.- 213 с.
16. Геоинформатика: учебник для студ. высш. учебн. заведений / Е.Г.Капралов, А.В.Кошкарев, В.С.Тикунов и др.; под ред. В.С.Тикунова.В 2 кн. Кн1.- 2-е изд. перераб. и доп. – М.: Издательский центр "Академия", 2008.- 384 с.
17. The GIS History Project [Electronic Resource] .- URL: http://www.ncgia.buffalo.edu/gishist/bar_harbor.html
18. Світличний О.О. Основи геоінформатики: Навчальний посібник / О.О.Світличний, С.В.Злозницький; за заг. ред.. О.О. Світличного. – Суми: ВТД Університетська книга, 2006. – 295 с.
19. Географічна інформація – Еталонна модель : Нац. стандарт України (ДСТУ ISO 19101:2002(E). – К.: Держспоживстандарт України, 2005. – 65 с.
20. David M. Mark. Geographic Information Science: Defining the Field [Electronic Resource] .- URL: <http://www.geog.buffalo.edu/~dmark/>
21. Huxhold W.E., Levinsohn A.G. Managing Geographic Information System Projects. - New York, Oxford: Oxford University Press., 1995. - 250 p.
22. The University Consortium for Geographic Information Science [Electronic Resource] .- URL: <http://www.ucgis.org/>
23. CSU Geospatial Review.- California State University GIS Specialty Center, 2006 [Electronic Resource] .- URL: <http://csugis.sfsu.edu/>
24. Савиных В.П. Геоинформационный анализ данных дистанционного зондирования / В.П.Савиных , В.Я. Цветков. - М.: Геодезиздат,- 2001 – 228 с.
25. David DiBiase, Michael DeMers, Ann Johnson, Karen Kemp, Ann Taylor Luck, Brandon Plewe, and Elizabeth Wentz. - Geographic Information Science & Technology. Body of Knowledge [Electronic Resource] .- URL: <http://www.ucgis.org/>
26. Берлянт А.М. Государственный образовательный стандарт по геоинформационным системам и проблемы деятельности Комитета ГИС-образование ГИС-Ассоциации [Electronic Resource] .- URL: http://loi.sccc.ru/gis/gisa/GIS_Educat/stand_eduac99.htm

27. Королев Ю.К. Общая геоинформатика. Часть 1. Теоретическая геоинформатика. Вып. 1./ Ю.К. Королев - М.: Дата+, 1998. - 118 с.
28. Understanding GIS - The ARC/INFO Method. - ESRI, 1990, 522 p.
29. Зейлер М. Моделирование нашего мира: Руководство ESRI по проектированию базы геоданных / М.Зейлер; пер. с англ. - М.: СП ООО Дата+, 2004. - 254 с.
30. Joseph K. Berry. Beyond Mapping III. Understanding Spatial Patterns and Relationships - BASIS Press, 2007, 227 p. [Electronic Resource] .- URL: <http://www.innovativegis.com/basis/MapAnalysis/>
31. Principles of Geographic Information Systems / Rolf A. de By (Ed.).- ITC, Enschede, The Netherlands. - 490 p.
32. GIS Glossary [Electronic Resource] .- URL: <http://www.geog.ubc.ca/courses/klink/gis.notes/glossary.html>
33. Spatial Analysis and GIS: A Primer / Gilberto Camara and other. - Image Processing Division, National Institute for Space Research (INPE), Brazil
34. Скворцов А.В. Триангуляция Делоне и её применение / А.В. Скворцов. – Томск: Изд-во Том. ун-та, 2002. – 128 с.
35. Класифікатор інформації, яка відображається на топографічних планах масштабів 1:5000, 1:2000, 1:1000, 1:500 / Наказ Головного управління геодезії, картографії і кадастра, при Кабінеті Міністрів України № 25 від 9.03.2000 р.
36. Техніко-економічна доповідь по формуванню національної інфраструктури геопросторових даних України (УкрНІГД) // Звіт НДІГК, 2005. - 112 с.
37. Томлинсон, Роджер. Думая о ГИС. Планирование географических информационных систем: Руководство для менеджеров / Томлинсон, Роджер. – М.: Изд. Дата+, 2004. - 329 с.
38. Иванников А.Д. Прикладная геоинформатика / А.Д. Иванников, В.П.Кулагин, А.Н. Тихонов, В.Я. Цветков. - М.: МАКС Пресс, 2005. - 360 с.
39. Средства подготовки данных [Electronic Resource] .- URL: <http://www.tsoring.ru/datapreapiment.htm>
40. Единая система классификации и кодирования. Термины и определения / ГОСТ 17369-85
41. Ситник В.Ф. Основи інформаційних систем: Навч. посібник / В.Ф.Ситник, Т.А.Писаревська, Н.В. Єрьоміна, О.С. Краєва ; за ред. В.Ф. Ситника. - К.: КНЕУ, 1997. - 252 с.
42. Шипулін В.Д. Створення базового набору геопросторових даних // Вчені записки ТНУ.- 2006.- Т. 19(58). - №2 – С. 151-156. (Серія Географія)

43. Melita Kennedy and Steve Kopp. Understanding Map Projections. - ESRI™, 2002. – 121 p.
44. World Geodetic System 1984, WGS 84 // TR 8350.2, Third Edition. – The National Imagery and Mapping Agency (NIMA), 2000.
45. Getting Started with Geographic Information Systems / Keith C. Clark - 4 ed. -Prentice Hall, 2003.- 342 p.
46. Harvey Francis. A Primer of GIS. - The Guilford Press, 2008.- 310 p.
47. Sarah Battersby and Nicholas Matzke.Lab 3: GIS Data Models - UC Santa Barbara, 2000 [Electronic Resource] .- URL: <http://cda.morris.umn.edu/~jonesjv/gis/labs/lab3.html>
48. David J. Buckley. The GIS Primer [Electronic Resource] .- URL: <http://www.innovativegis.com/education/primer.html>
49. Объектно-ориентированная методология [Electronic Resource] .- URL: <http://belani.narod.ru/3/OOM.htm>
50. ArcGIS Desktop Help [Electronic Resource] .- URL: <http://webhelp.esri.com/arcgisdesktop/9.3/>
51. Энди Митчелл. Руководство по ГИС Аналізу. Часть 1: Пространственные модели и взаимосвязи / Энди Митчелл; пер. с англ. – Киев, ЗАО ECOMM Co; СтилоС, 2000. – 198 с.
52. Ішук О. О. Просторовий аналіз і моделювання в ГІС: Навчальний посібник / О. О.Ішук, М. М.Коржнев, О. Е.Кошляков; за ред. акад. Д.М.Гродзинського. – К.: Видавничо-поліграфічний центр "Київський університет", 2003.– 200 с.
53. Aronoff Stan. Geographic Information Systems: A Management Perspective. WDL Publications, Ottawa, Canada, 1989.
54. Principles of Geographic Information Systems. Rolf A. de By (ed.). Second edition.– Enschede, The Netherlands, 2001, 490 p.
55. Berry, J. K. 1989. Beyond mapping: Spatial data analysis. pp. 10–16. In The GIS Sourcebook. GIS World, Inc., Ft. Collins, CO. 180 p.
56. Chrisman Nicholas. Exploring Geographical Information Systems. 2 edition John Wiley & Sons – 2003, 306 p.
57. ArcGIS 9. Using ArcMap.– ESRI, 2005, 598 p.
58. GIS self learning tool [Electronic Resource] .- URL: <http://www.sli.unimelb.edu.au/gisweb/>
59. Geospatial Analysis - a comprehensive guide. 2nd edition © 2006-2008 de Smith, Goodchild, Longley

Рекомендуемая литература для самостоятельного изучения

Основная литература

1. Геоинформатика: учебник для студ. высш. учебн. заведений / Е.Г.Капралов, А.В.Кошкарев, В.С.Тикунов и др.; под ред. В.С.Тикунова. В 2 кн. Кн1.- 2-е изд. перераб. и доп. – М.: Издательский центр "Академия", 2008.- 384 с.
2. ДеМерс М.Н. Географические информационные системы. Основы / ДеМерс М.Н.; пер. с англ. - М.: Дата+, 1999.- 491 с.
3. Світличний О.О. Основи геоінформатики: Навчальний посібник / О.О.Світличний, С.В.Злотницький; за заг. ред. О.О. Світличного. – Суми: ВТД Університетська книга, 2006. – 295 с.

Дополнительная литература

1. Журкин И. Г. Геоинформационные системы / И. Г. Журкин, С. В.Шайтура. - М : КУДИЦ-ПРЕСС, 2009. - 272 с.
2. Зейлер М. Моделирование нашего мира: Руководство ESRI по проектированию базы геоданных / М.Зейлер; пер. с англ. - М.: СП ООО Дата+, 2004. - 254 с.
3. Ішук О. О. Просторовий аналіз і моделювання в ГІС: Навчальний посібник / О. О.Ішук, М. М.Коржнев, О. Е.Кошляков; за ред. акад. Д.М.Гродзинського. – К.: Видавничо-поліграфічний центр "Київський університет", 2003.– 200 с.
4. Королев Ю.К. Общая геоинформатика. Часть 1. Теоретическая геоинформатика. Вып. 1./ Ю.К. Королев - М.: Дата+, 1998. - 118 с.
5. Лурье И. К. Геоинформационное картографирование. Методы геоинформатики и цифровой обработки космических снимков: учебник / И.К.Лурье. - М.: КДУ, 2008, - 424 с.: с илл.,табл.
6. Митчелл Энди. Руководство по ГИС Аналізу. Часть 1: Пространственные модели и взаимосвязи / Митчелл Энди; пер. с англ. – Киев, ЗАО ЕСОММ Со; Стилос, 2000. – 198 с.
7. Иванников А.Д. Прикладная геоинформатика / А.Д. Иванников, В.П.Кулагин, А.Н. Тихонов, В.Я. Цветков. - М.: МАКС Пресс, 2005. - 360 с.
8. Хаксхольд Виллиам. Введение в городские геоинформационные системы / Хаксхольд Виллиам; пер. с англ. - М.: Дата+, 1998.- 321 с.

Навчальне видання

Володимир Дмитрович Шипулін

**ОСНОВНІ ПРИНЦИПИ ГЕОІНФОРМАЦІЙНИХ
СИСТЕМ**

**Навчальний посібник
(Рос. Мовою)**

Редактор *М. З. Аляб'єв*

Комп'ютерне верстання *С. Г. Панова*

Дизайн обкладинки *Т. Є. Клочко*

Підп. до друку 22.12.09 р.
Друк на ризографі
Тираж 500 пр.

Формат 60x84 1/16
Ум. друк.арк. 14,04
Зам. №

Видавець і виготовлювач:

Харківська національна академія міського господарства,
вул. Революції, 12, Харків, 61002

Електронна адреса: rectorat@ksame.kharkov.ua

Свідоцтво суб'єкта видавничої справи: ДК №731 від 19.12.2001