Ученые записки Таврического национального университета им. В. И. Вернадского
Серия «ГЕОГРАФИЯ» Том 19 (58) № 2 (2006) 119-131

*УДК 004.9:004.65*

# ПРИМЕР ПОДХОДА К ФОРМИРОВАНИЮ СТРУКТУРЫ НАЦИОНАЛЬНЫХ ПРОСТРАНСТВЕННЫХ ДАННЫХ УКРАИНЫ

***Салтовец А.А., Николаев В.М., Ломоносова О.С.***

Национальные инфраструктуры пространственных данных (NSDI) являются в настоящее время одной из наиболее быстро развивающихся областей мирового информационного пространства. Определяющими факторами такого развития являются совершенствование программных средств ГИС и WEB-технологий, растущая доступность и детальность космических снимков, внедрение в повседневную жизнь систем спутникового позиционирования. Эти факторы формируют постоянно растущий спрос на разнообразные сервисы предоставления пространственных данных.

Обобщенно говоря, инфраструктура пространственных данных включает в себя [1] следующие составляющие (Рис. 1):

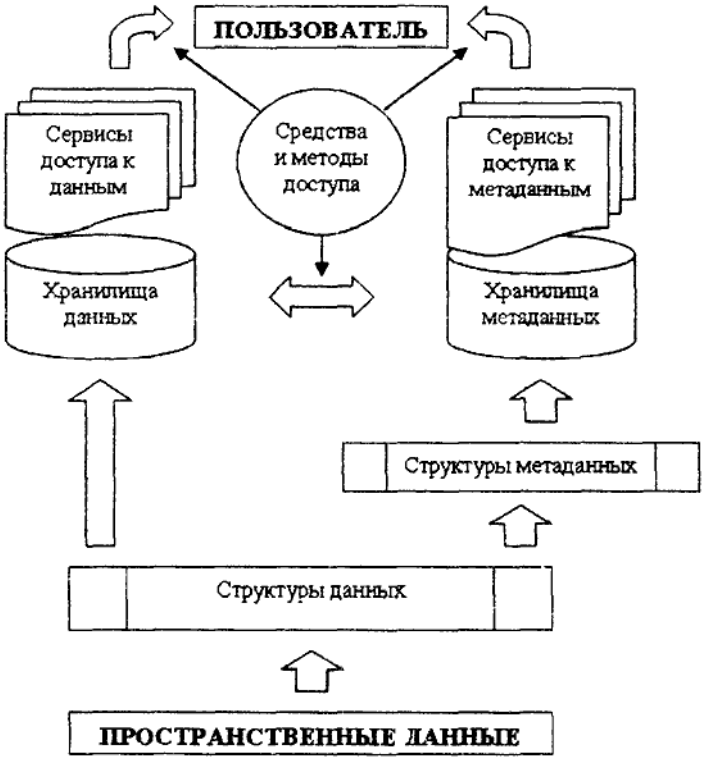


Рис. 1. Обобщенная схема инфраструктуры пространственных данных

В этой статье хотелось бы акцентировать внимание на собственно структуре пространственных данных, как фундаменте всей инфраструктуры. Отнюдь не умаляя значения других составляющих инфраструктуры, следует подчеркнуть, что без создания структуры данных создание других компонентов инфраструктуры становится бессмысленным.

Сегодня мы имеем перед собой для изучения и сравнения целый ряд более и менее успешных проектов структур данных, которые используются в различных национальных и наднациональных SDI. Это, в первую очередь, Соединенные Штаты, Япония, Канада, Австралия, Единая Европа и др.[2]. В последние два года концепция NSDI разрабатывается и в России – ГИС-Ассоциацией и Министерством транспорта и связи России.

Рассматривая пути и методы разработки структур пространственных данных в различных странах, можно сделать наиболее важный методологический вывод – чем меньше жестких стандартов на форматы данных, объектный состав, атрибутивные справочники и т.п. и чем больше стандартов, описывающих принципы построения и описания данных, тем быстрее и с меньшими затратами такие структуры внедряются в повседневную жизнь страны.

Одной из первых на территории СНГ попыток упорядочивания пространственных данных была разработанная нами в содружестве со специалистами военной и гражданской картографии система нормативов по составлению и использованию электронных картматериалов. Эти нормативы были утверждены Укргеодезкартографией в 1998 г. [3] и достаточно широко распространены как среди разработчиков, так и среди пользователей пространственных данных в Украине, России, Белоруссии и Молдове. Они сыграли свою положительную роль в упорядочении пространственных данных при решении многих практических задач.

Однако, по мере продвижения геоинформационных технологий в сферы государственного управления, военных задач, крупного бизнеса, мы стали наблюдать, что жесткая структура классификатора и атрибутивного состава становятся тормозом для широкого распространения пространственных данных. Более того, использование единой жесткой структуры, инвариантной к типу решаемых задач, зачастую значительно усложняет разработку прикладных информационных систем.

Основным принципом построения нормативов была преемственность по отношению к топографической карте и использование в модифицированном виде топографического классификатора с соответствующим атрибутивным составом.

Отсутствие при таком подходе отношений (связей, наследования) между сущностями и стремление сохранить свойственную топографической карте атрибутику приводит к значительному дублированию информации, делает структуру громоздкой и усложняет разработку ПО. Особенно ярко это проявляется при использовании в одной информационной системе разномасштабных пространственных данных.

Сегодня мы приходим к выводу о необходимости использования более гибкого подхода к проектированию структуры данных. Этот подход базируется на двух

основных принципах – данные должны агрегироваться по признакам своего использования, а структура каждого сегмента должна строиться как модель реального мира с использованием объектно-ориентированной методологии.

При этом представление сущностей в модели должно быть инвариантно по отношению к масштабу, а атрибутивный состав минимизирован с использованием общих справочников.

В настоящей работе для описания моделей используется язык UML и набор классов ArcGIS от фирмы ESRI. Естественно, описание модели средствами UML позволяет реализовать модель на любом другом наборе классов ГИС.

Проиллюстрируем этот подход, построив в качестве одного из сегментов NSDI сегмент «Административно-территориальное деление и населенные пункты». В состав этого сегмента включим единицы административно-территориального деления, населенные пункты, справочник КОАТУУ и данные переписи населения. Как видно, эти данные сгруппированы, исходя из единства целей и методов их использования – административное управление и планирование, демографический анализ, маркетинг, обслуживание населения и т.п.

В традиционном подходе топографического классификатора территориально-административное деление относилось к сегменту данных «Политико-административное устройство, ограждения, отдельные природные явления и объекты», а населенные пункты составляли отдельный одноименный сегмент. Данные же переписи населения содержались в виде различных атрибутов в обоих сегментах. На Рис. 2 и Рис. 3 приведены UML-диаграммы структуры указанных сегментов в соответствии с топографическим классификатором.

Рассмотрение этих диаграмм показывает, что применение такой структуры для создания «базового» или «опорного», с точки зрения единства использования, сегмента совершенно нерационально.

Во-первых, в структуре присутствуют сущности, которые являются лишними для сформулированных выше целей – это всякого рода природные явления (тропики, аномалии магнитного склонения), национальные парки и заповедники, разнообразные ограждения и т.п.

Во-вторых, при попытке разместить в одном хранилище пространственные данные различной детальности (масштаба), необходимо будет организовать отдельные наборы данных для каждого масштаба. При этом внутренняя структура у каждого такого набора будет иметь как общие свойства, так и значительные отличия. Это приведет, с одной стороны, к дублированию данных, а с другой стороны – к отсутствию единства их представления.

В-третьих, такие существенные с точки зрения использования данные, как численность населения и коды КОАТУУ, представлены в структуре в виде отдельных атрибутов различных классов, что не дает возможности использования в целостном виде таких общегосударственных справочников, как справочник КОАТУУ и статистические справочники по населению. Кроме того, ошибки, вносимые в процессе актуализации несвязанных атрибутов у различных классов, будут постоянно нарушать целостность данных.

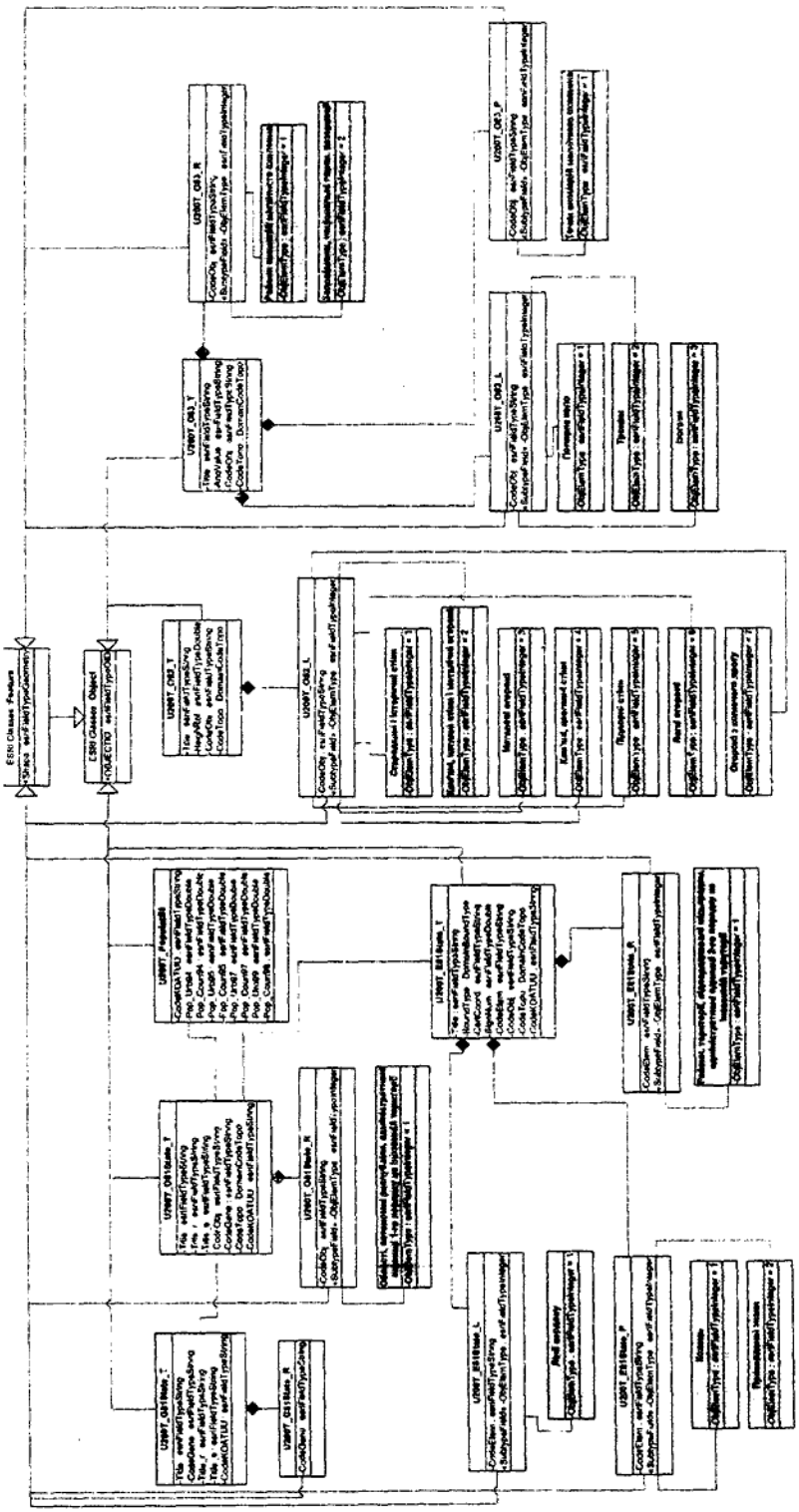

Рис. 2 Сегмент «Политико-административное устройство...»

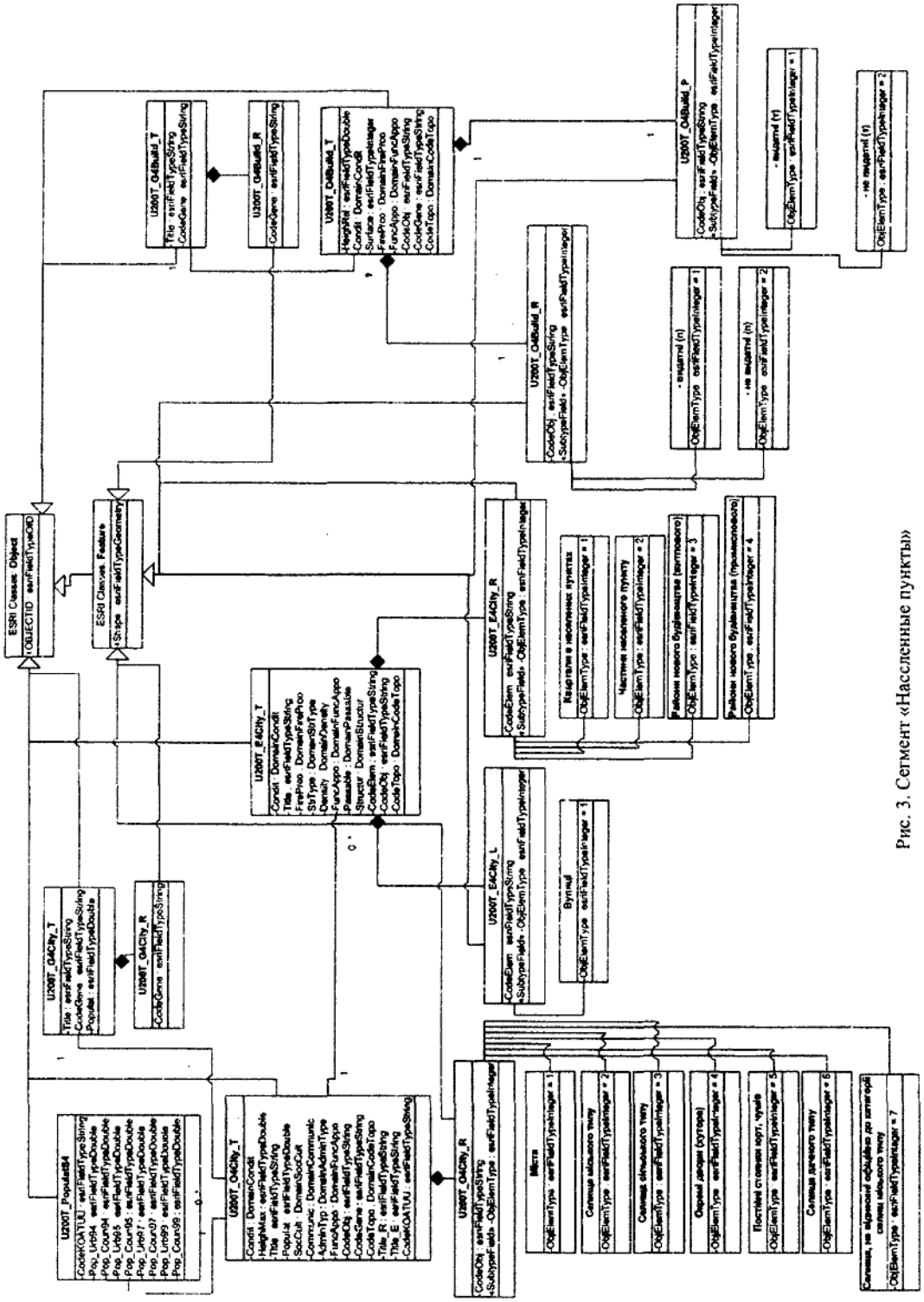


Рис. 3. Сегмент «Населенные пункты»

В-четвертых, атрибутивный состав топографического классификатора с точки зрения использования такого сегмента как «базового», в значительной степени избыточен. Большое количество атрибутов могут быть отнесены к «целевым» сегментам NSDI, создаваемым для решения частных прикладных задач.

Рассмотрим структуру создаваемого сегмента. Для удобства представления структура отражена пятью UML-диаграммами. На первой диаграмме (Рис. 4) представлено территориальное деление Украины, которое включает в себя три уровня территориальных образований – государство, область, район.

Соответственно, создано три абстрактных класса пространственных объектов с геометрическим типом «полигон» – *U_State*, *U_Reg* и *U_SReg*. Три класса необходимы для обеспечения различных систем кодификации этих пространственных объектов.

Классами-реализациями являются собственно классы пространственных объектов этих образований. Каждый абстрактный класс порождает по несколько потомков, так как контуры полигонов территориальных образований имеют различную детальность для различных масштабов. В данной модели организовано по три класса, исходя из допущения о разбивке масштабного ряда на три диапазона 1:4 000 000, 1:1 000 000, 1:500 000 – 1:200 000.

Пространственные классы не имеют никакой атрибутики кроме уникального идентификатора. Все необходимые характеристики помещены в атрибутивные объектные классы, с которыми пространственные классы связаны отношениями. Соответственно, все три пространственных класса связаны только с одним объектным классом, что обеспечивает единство представления данных для любого масштаба отображения.

Для сохранения целостности данных эти отношения являются композитными связями, что обеспечивает каскадное удаление сущностей в атрибутивном и пространственном классах средствами СУБД.

В свою очередь, набор объектных классов административно-территориального деления состоит из класса КОАТУУ и абстрактного класса единиц административного деления, потомками которого являются классы государств, областей, районов и советов (Рис. 5). Эти классы (за исключением класса государств) имеют подтипы, соответствующие реальной административной структуре Украины.

Пространственные классы населенных пунктов (аналогично пространственным классам территориального деления) представлены тремя классами для отображения разномасштабных данных (Рис. 6). Следует отметить, что в отличие от классов территориального деления, классы населенных пунктов имеют как полигональный геометрический тип, так и точечный для мелких масштабов.

Все три пространственных класса связаны композитными отношениями с единственным объектным классом, содержащим характеристики населенных пунктов.

Как и в случае административно-территориального деления, подтипы при этом организованы в объектных, а не в пространственных классах, для соблюдения единства представления разномасштабных данных (Рис. 7.)

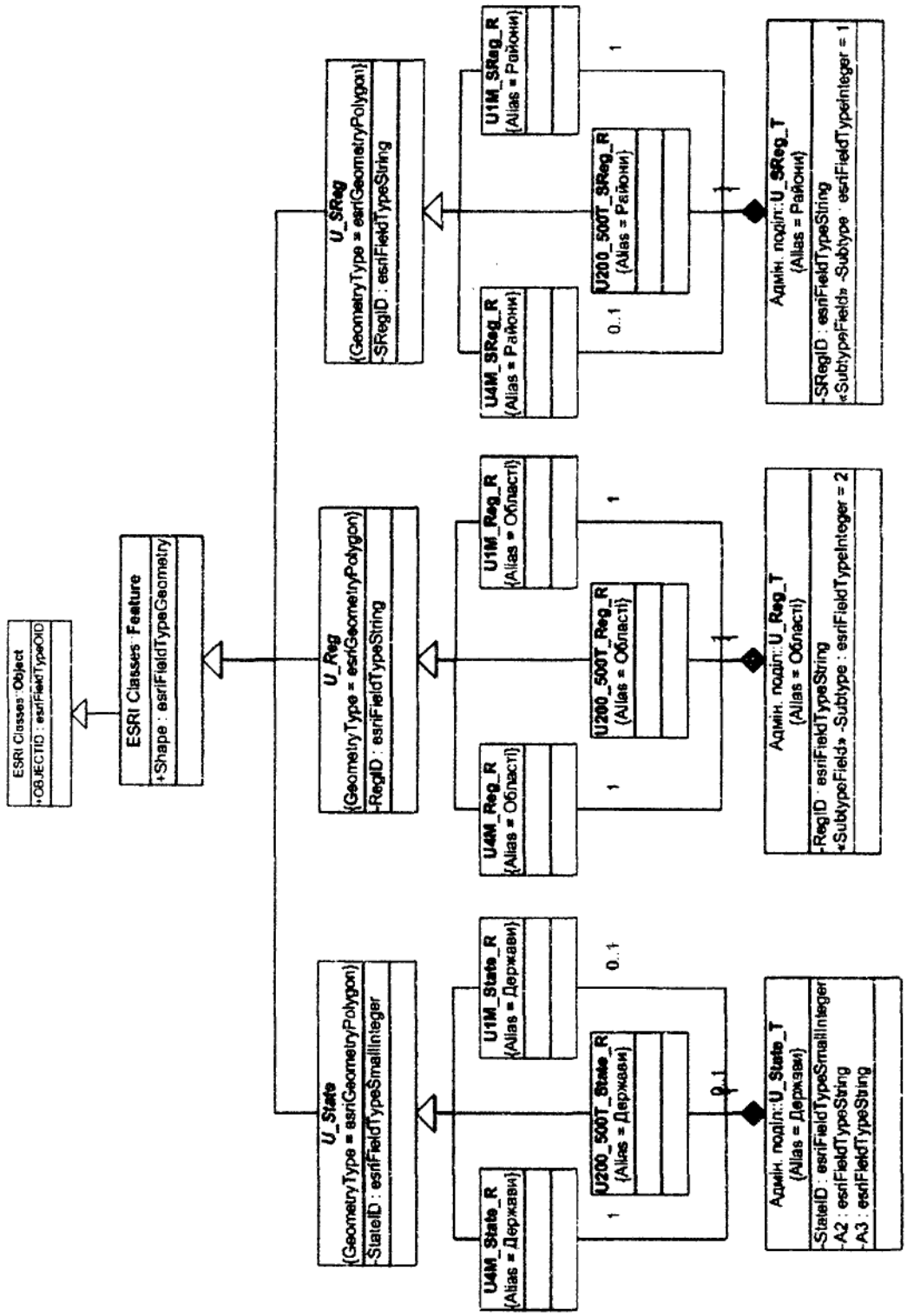


Рис. 4. Территориальное деление

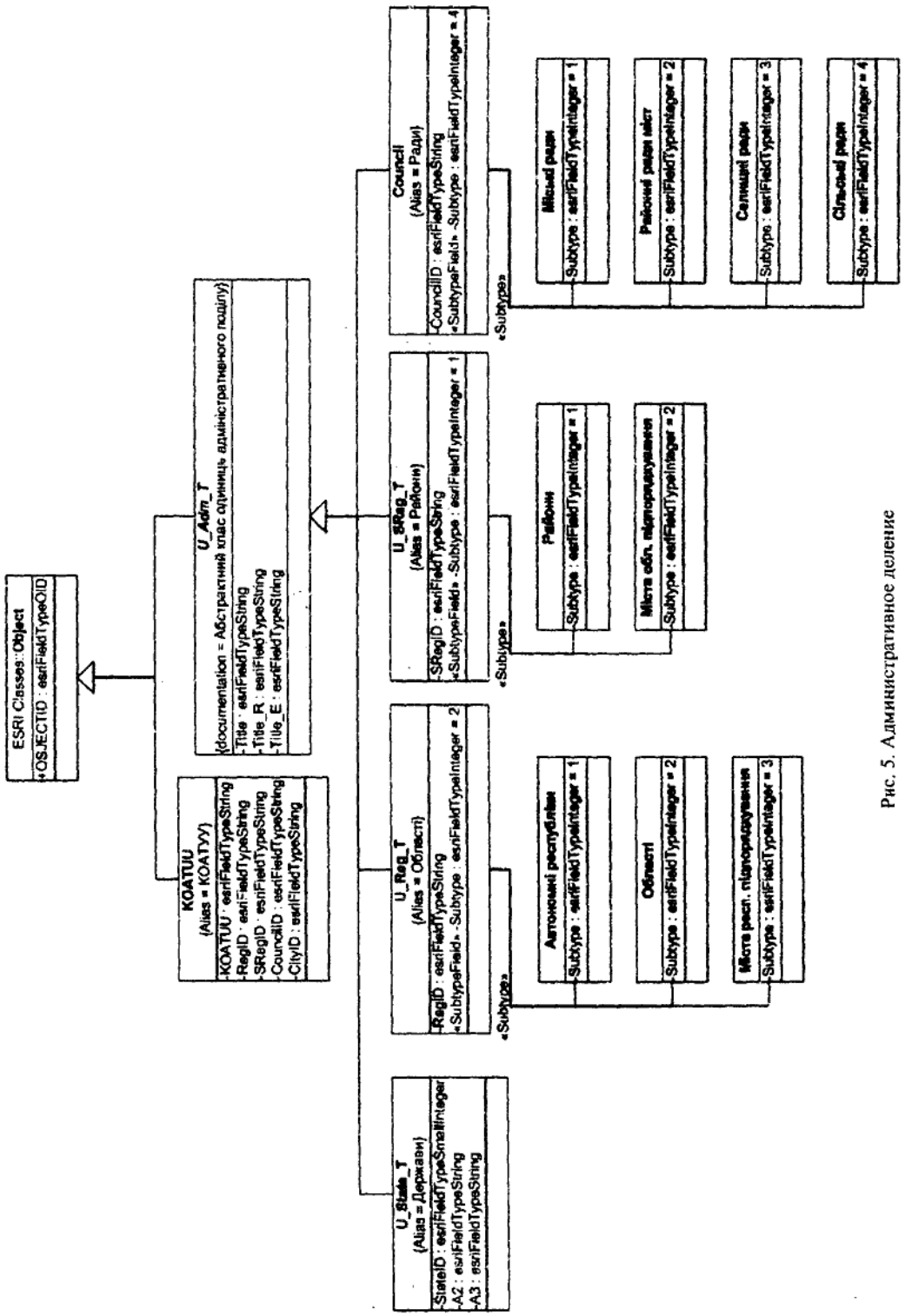


Рис. 5. Административное деление

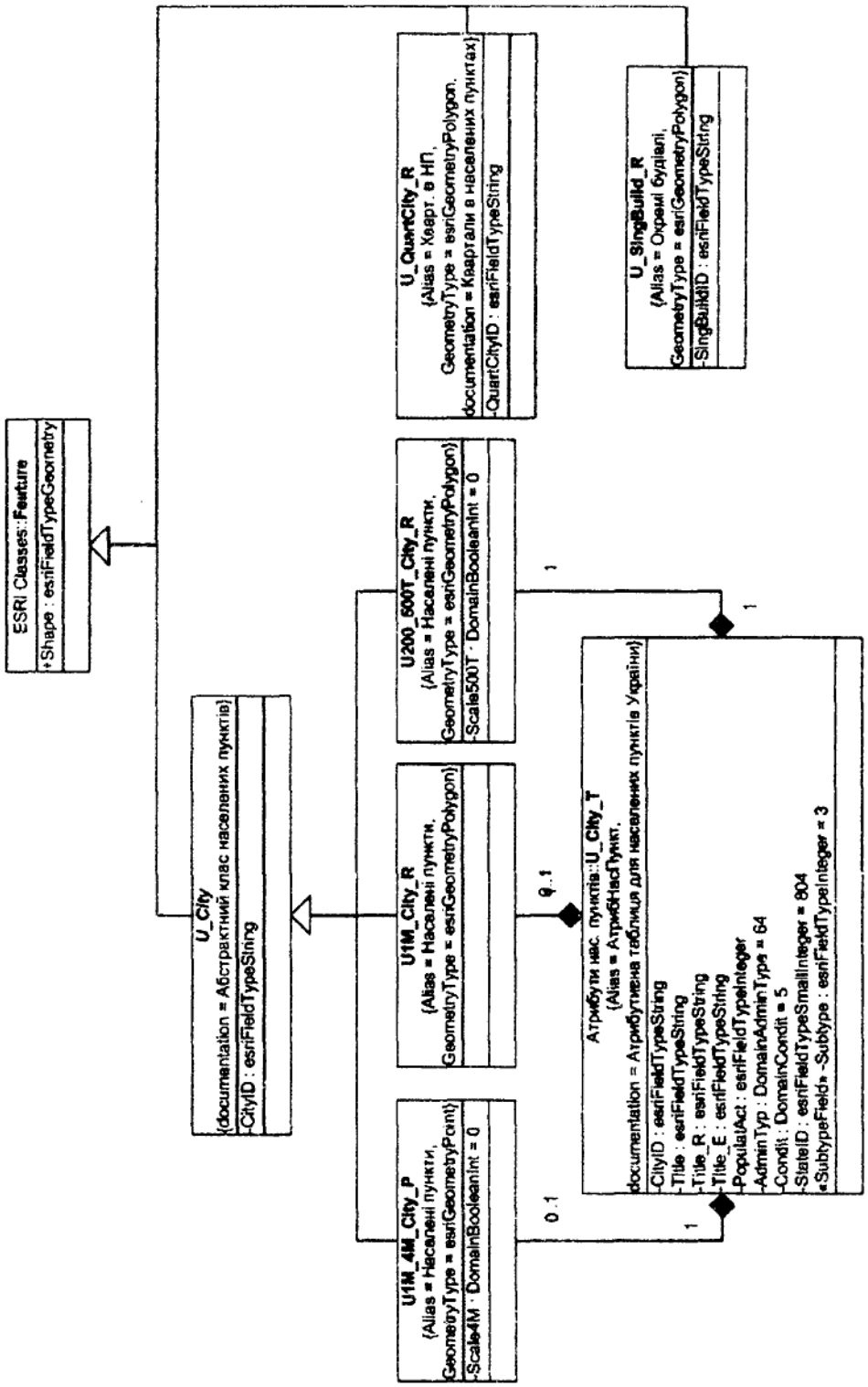


Рис. 6. Населенные пункты

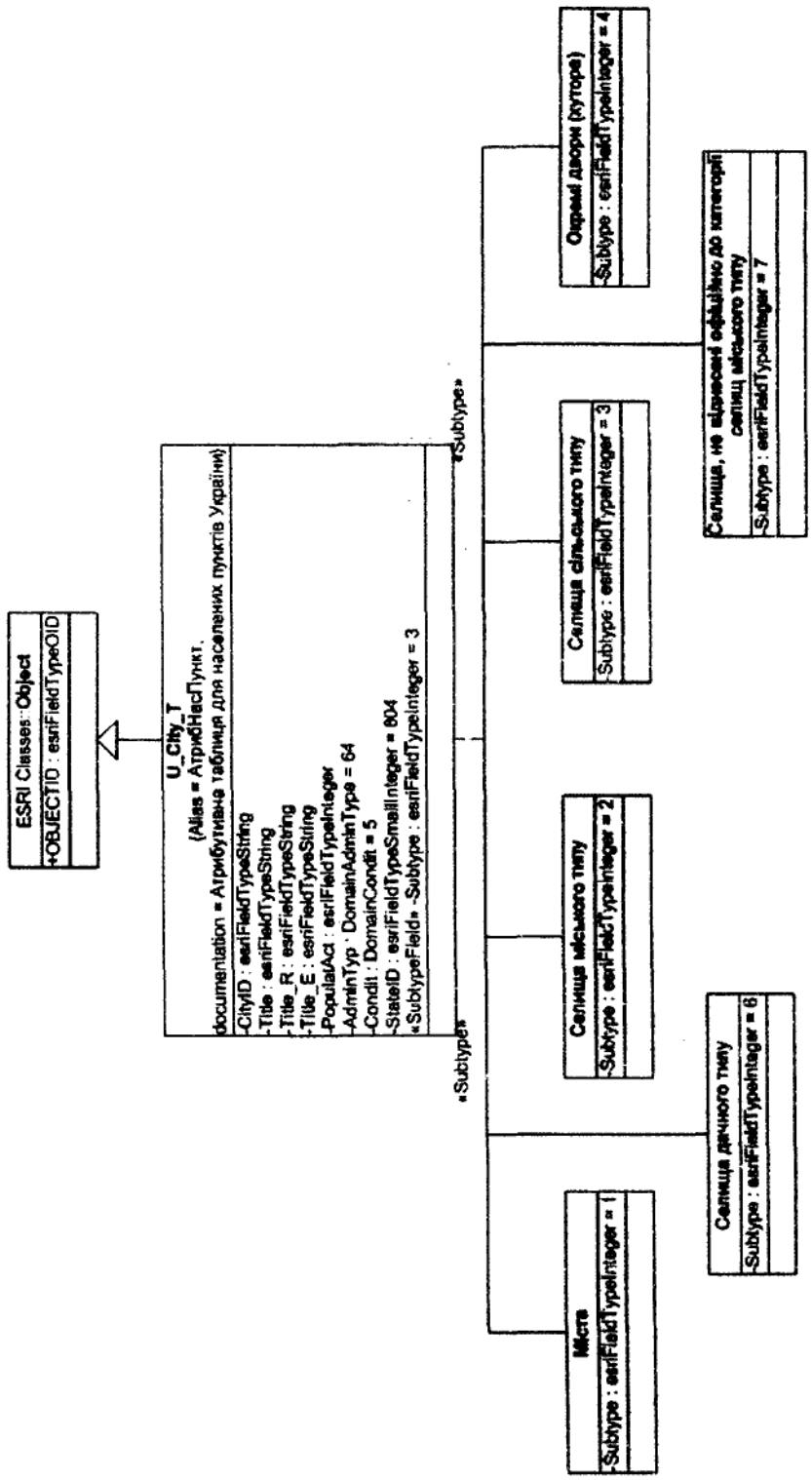


Рис. 7. Объектные классы населенных пунктов

Наиболее существенной диаграммой, иллюстрирующей внутренние связи сегмента, является диаграмма, приведенная на Рис. 8. Все внутренние связи сегмента реализованы с помощью объектных классов, что делает структуру инвариантной к выбранной ГИС-технологии.

Справочник КОАТУУ является центральным объектным классом диаграммы, обеспечивая связь населенных пунктов и административно-территориального деления. Справочник населения представлен двумя объектными классами, одни из которых содержит сведения о населении административно-территориальных единиц, а второй – о населении населенных пунктов.

Следует также обратить внимание на кардинальность связи между классом населенных пунктов U_City_T и классом КОАТУУ – 1 к 0...1. Такая кардинальность отражает ситуацию, когда населенный пункт как физический объект присутствует на местности, но является разрушенным или отселенным и, вследствие этого, не имеет кода КОАТУУ.

Построенная таким образом структура сегмента отвечает всем требованиям к опорным данным NSDI и, естественно, может быть легко дополнена необходимыми атрибутами.

Эта структура была реализована в Правительственной Информационно-аналитической системе по чрезвычайным ситуациям как часть геобазы данных ArcGIS 9 на СУБД MS SQLServer с использованием ArcSDE. Планируется разработка на этих же принципах других сегментов центральной геобазы.

В заключение хотелось бы также отметить, что, несмотря на значительную гибкость самих логических структур данных, две составляющие стандартизации NSDI должны быть относительно жесткими.

Это вопросы стандартизации описания метаданных и принятие национальной системы уникальных кодификаторов данных. И если по вопросу метаданных в мире уже разработано несколько как национальных, так и международных стандартов [4,5], на которые легко опираться, то система кодификации слишком индивидуальна для каждой страны, связана с уже имеющими системами уникального кодирования и должна строиться очень продуманно – ведь любое изменение однажды присвоенных кодов повлечет за собой лавинообразный эффект перестроек баз данных и переделок программного обеспечения.

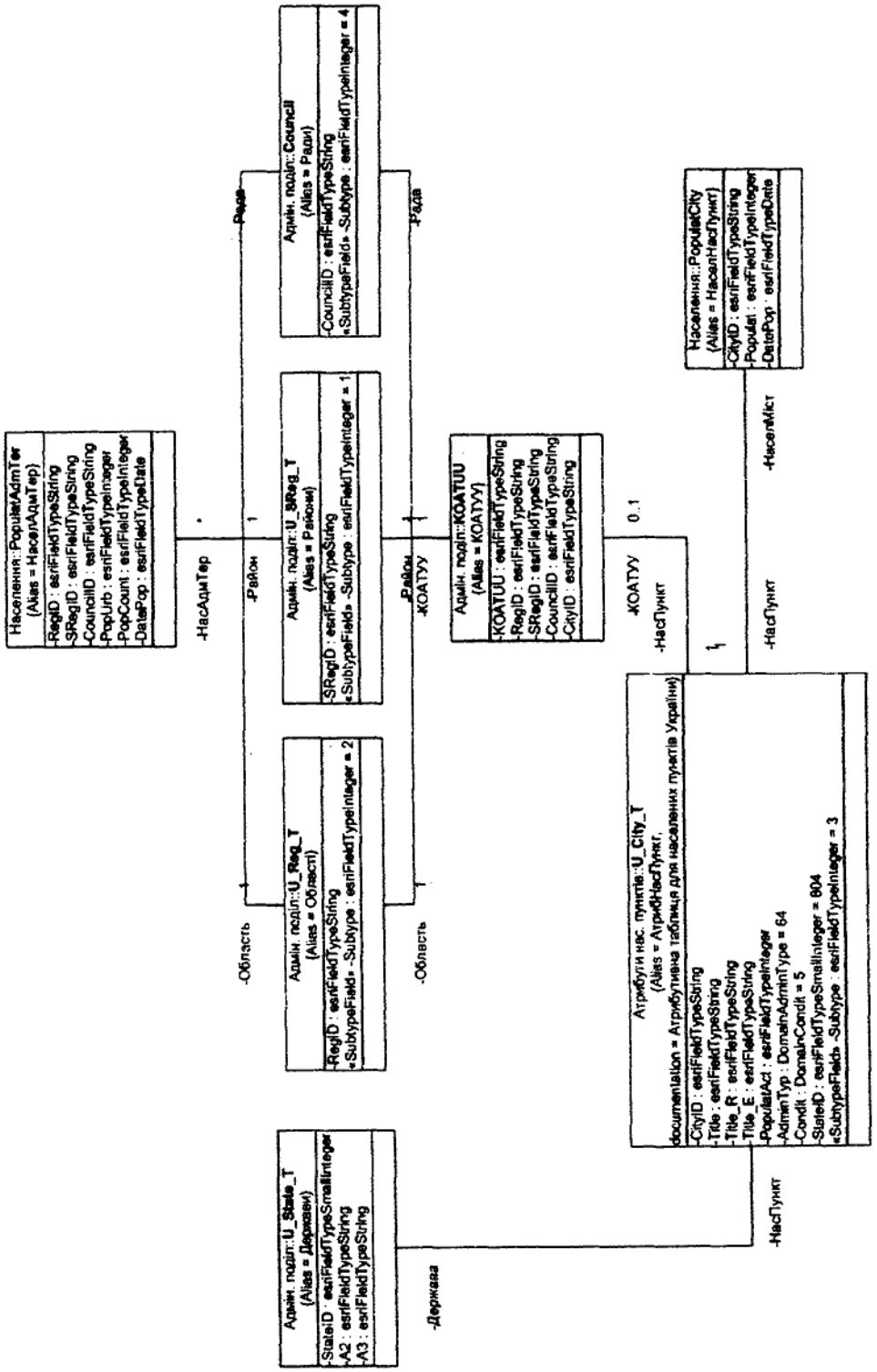
Населення::PopulatAdmTer
(Alias = НаселАдмТер)
-RegID : esriFieldTypeString
-SRegID : esriFieldTypeString
-CouncilID : esriFieldTypeString
-PopUrb : esriFieldTypeInteger
-PopCount : esriFieldTypeInteger
-DataPop : esriFieldTypeDate
-НаселАдмТер
*
Адмін. поділ::U_Reg_T
(Alias = Області)
-RegID : esriFieldTypeString
«SubtypeField» -Subtype : esriFieldTypeInteger = 2
-Область
1
Адмін. поділ::U_SReg_T
(Alias = Райони)
-SRegID : esriFieldTypeString
«SubtypeField» -Subtype : esriFieldTypeInteger = 1
-Район
1
Адмін. поділ::Council
(Alias = Ради)
-CouncilID : esriFieldTypeString
«SubtypeField» -Subtype : esriFieldTypeInteger = 4
-Рада
-Рада
-Область
-Район
-КОАТУУ
Адмін. поділ::KOATUU
(Alias = КОАТУУ)
-KOATUU : esriFieldTypeString
-RegID : esriFieldTypeString
-SRegID : esriFieldTypeString
-CouncilID : esriFieldTypeString
-CityID : esriFieldTypeString
0..1
-КОАТУУ
-НаселПункт
Населення::PopulatCity
(Alias = НаселНаселПункт)
-CityID : esriFieldTypeString
-Populat : esriFieldTypeInteger
-DataPop : esriFieldTypeDate
-НаселМіст
-НаселПункт
1
Атрибути нас. пунктів::U_City_T
(Alias = АтрибНаселПункт.
documentation = Атрибутивна таблиця для населених пунктів України)
-CityID : esriFieldTypeString
-Title : esriFieldTypeString
-Title_R : esriFieldTypeString
-Title_E : esriFieldTypeString
-PopulatAct : esriFieldTypeInteger
-AdminTyp : DomainAdminType = 64
-Condit : DomainCondit = 5
-StateID : esriFieldTypeSmallInteger = 804
«SubtypeField» -Subtype : esriFieldTypeInteger = 3
-НаселПункт
Адмін. поділ::U_State_T
(Alias = Держави)
-StateID : esriFieldTypeSmallInteger
-A2 : esriFieldTypeString
-A3 : esriFieldTypeString
-Держава

Рис. 8. Связи

## Список литературы

1. Кошкарев А.В. Инфраструктуры пространственных данных // ГИС-обозрение. - 2000. - № 3-4. .- С. 5 - 10 (начало); - 2001. - № 1. - С. 28 – 32
2. Карпінський Ю.О., Лященко АА., Волчко Є.П. Стандартізація географічиої інформації. Міжнародний досвід та шляхи розвитку в Україні // Вісник геодезії та картографії. - 2002. - №3(26) - С. 32-38.
3. Нормативи по створенню електронних карт місцевості // Укргеодезкартографія, 1998, С. 186
4. Standards for Digital Geospatial Metadata. Federal Geographic Data Committee. June 8, 1994.-54 pp.
5. Standards for Spatial Metadata. Federal Geographic Data Committee. Draft, November 3, 1992.J)28 pp.

Статья поступила в редакцию 03.05.06