

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ

Государственное образовательное учреждение высшего профессионального образования

САМАРСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

УДК _____

Л.А. Лыноградский, А.Г. Мищенко, Д.А. Писарев

КЛАССИФИКАЦИЯ КОНЦЕПЦИЙ ЭКОЛОГИЧЕСКОГО РАЗВИТИЯ
РЕГИОНОВ НА ОСНОВЕ КЛАСТЕРНОГО АНАЛИЗА

Самара 2010

Содержание

Введение.....	3
Методика кластеризации.....	6
1. Выделение из текста существенных признаков.....	6
2. Группировка задач.....	12
3. Построение таблицы суммарных весов задач объектов по группам.....	17
4. Построение матрицы расстояний объектов.....	18
Результаты и их обоснование.....	21
Итоги.....	26

Введение

Устойчивое социально-экономическое развитие региона невозможно без обеспечения высоких экологических стандартов окружающей среды, сохранения биологического разнообразия, эффективного и рационального использования природных ресурсов территории. Достижение указанных параметров является основой повышения уровня жизни и обеспечения экологической безопасности и экологических интересов населения и долгосрочной целью государственного управления в сфере природопользования и охраны окружающей среды. Высокая степень хозяйственной освоенности территории, развитие производства и концентрация населения привели к значительному антропогенному воздействию и изменениям в природе, что придает особую актуальность разработке и реализации эффективной региональной политики и государственного управления в сфере природопользования и охраны окружающей среды.

Разработка концепции экологического развития призвана определить основные проблемы, задачи, приоритеты, методы, этапы и механизмы решения экологических проблем в регионе как неотъемлемой составляющей его социально-экономического развития.

Авторы провели поиск документов, представленных на официальных сайтах региональных министерств, других учреждений, а также в СМИ, которые могут рассматриваться как базовые экологические концепции:

- концепции экологической политики,
- концепции экологической безопасности,
- концепции экологического развития
- целевые программы по экологической безопасности
- целевые программы по охране окружающей среды,
- целевые программы «Экология и природные ресурсы».

По данным, размещенным в сети Интернет, полнотекстовые экологические концепции имеют 12 регионов, упоминание о существовании экологической концепции региона с указанием основных направлений экологической политики присутствует в 1 регионе, перечень мероприятий по целевым программам в области экологии имеется у 21 региона, сообщение о наличии экологической концепции в регионе без указания

основных направлений имеется еще у 3 регионов. В целом, упоминание о наличии экологических концепций или программ было найдено в 43 регионах РФ из 90.

Можно с уверенностью сказать, что актуальность разработки концепций регионов постепенно возрастает. Даже в тех регионах, где она уже разработана, ее качество не является очевидным, а изменения в состоянии экологической ситуации потребуют вносить количественные и качественные коррективы. Экологическая концепция региона имеет, как правило, следующие структурные элементы:

1. Цели и задачи концепции
2. Общие положения
3. Характеристика существующих экологических проблем региона
4. Основные направления реализации концепции
5. Целевые показатели, индикаторы, оценка ожидаемой эффективности, принципы осуществления
6. Механизмы реализации концепции, пути и средства реализации
7. Обоснование основных направлений концепции
8. Этапы реализации концепции
9. Основные понятия, глоссарий терминов
10. Разделы, посвященные финансовому обеспечению концепции, правовым вопросам, предполагаемым исполнителям и т.п.

На сегодняшний день нет разработанных методологических основ по формированию экологической концепции, нет устоявшегося подхода к оценке экологического состояния отдельной территории. Данная ситуация обусловлена географическими и экономическими отличиями различных регионов, которые влияют на экологическое состояние и цели экологического развития. Указанные различия также затрудняют обобщение опыта формирования экологической концепции.

Вместе с тем, на практике специалисты ориентированы на использование уже готовых концепций других регионов в качестве прототипов. Речь идет не только о копировании текста концепции, но прежде всего о возможности использования опыта в формировании организационно-административной структуры, законодательных документов, использовании критериев и моделей, методов и средств решения программных задач.

Сказанное позволяет сделать вывод — при всей индивидуальности регионов существуют типовые решения, предпочтительные для одной группы областей и малоэффективные для другой. Такого рода классификация лучше всего может быть

представлена в виде кластера, отражающего близость тех или иных регионов. Однако для построения кластера необходимо позиционировать объекты в пространстве и рассчитывать расстояние между ними, то есть пользоваться количественными характеристиками. Сравнение текстовых документов, какими являются концепции, затрудняет проведение кластерного анализа в его традиционном виде.

Авторы в настоящем исследовании предлагают подход, позволяющий проводить кластеризацию объектов, имеющих текстовое описание. На данном этапе подход реализован в применении к экологическим концепциям.

I. Методика кластеризации

1. Выделение из текста существенных признаков

На данном этапе необходимо выбрать признак, по которому можно обоснованно определить близость документов. Выбираемый признак, во-первых, должен обладать высокой информативностью; во-вторых, - присутствовать в явном виде во всех документах. В рассматриваемой области удобно использовать задачи и направления экологической политики. В большинстве случаев они имеют четкую формулировку, встречаются во всех документах.

В силу того, что не у всех субъектов РФ был найден такой документ как экологическая концепция (концепция экологической безопасности, концепция экологической политики), то были взяты в расчет и другие стратегические экологические документы, а также краткие обзоры документов, размещенные на интернет ресурсах, в которых четко выделены задачи концепции/доктрины/целевой программы. Найденные отдельные подпрограммы, например, подпрограмма «Леса», размещенные без основного документа (программы), не попали в список анализируемых объектов.

Таблица 1. Объекты кластеризации

№	Код региона	Название региона	Наименование документа
1	3	Республика Бурятия	Концепция экологизации экономики республики Бурятия
2	5	Республика Дагестан	Решение «Законодательное обеспечение охраны окружающей среды и экологической безопасности Республики Дагестан»
3	10	Республика Карелия	Республиканская целевая программа "Экология и природные ресурсы Республики Карелия на 2004-2010 годы"
4	14	Саха-Якутия	Концепция экологической безопасности республики Саха (Якутия)
5	16	Республика Татарстан	Концепция экологической безопасности Республики Татарстан (на 2007 - 2015 годы)
6	17	Республика Тува	Концепция экологической безопасности

			Республики Тува (до 2011 г.)
7	21	Чувашская республика	Республиканская целевая программа «Повышение экологической безопасности в Чувашской Республике на 2006-2010 годы»
8	26	Ставропольский край	Ведомственная целевая программа «Экология и природные ресурсы Ставропольского края на 2009-2011 годы»
9	29	Архангельская область	Концепция экологической политики Архангельской области на 2002–2010 годы
10	30	Астраханская область	Концепция комплексной целевой программы "Охрана окружающей среды Астраханской области на 2008 - 2010 годы"
11	35	Вологодская область	Ведомственная целевая программа департамента природных ресурсов и охраны окружающей среды Вологодской области "Экологическая безопасность и рациональное природопользование Вологодской области на 2009 - 2011 годы"
12	36	Воронежская область	Областная целевая программа «Экология и природные ресурсы Воронежской области на 2010-2014 годы»
13	41	Камчатская область	Долгосрочная краевая целевая программа "Охрана окружающей среды и обеспечение экологической безопасности в Камчатском крае на 2010-2012 годы"
14	43	Кировская область	Областная целевая программа «Экология и природные ресурсы Кировской области»
15	44	Костромская область	Постановление Костромской областной Думы от 30 октября 2008 г. N 3346 "Об утверждении областной целевой программы "Отходы" на 2009-2013 годы"
16	45	Курганская область	Ведомственная целевая программа Департамента природных ресурсов и охраны окружающей среды Курганской области «Охрана окружающей среды и экологическая

			безопасность Курганской области на 2008–2010 годы»
17	46	Курская область	Областная целевая программа «Экология и природные ресурсы Курской области (2009-2010 годы)»
18	51	Мурманская область	Ведомственная целевая программа Министерства природных ресурсов и экологии Мурманской области «Охрана и гигиена окружающей среды и обеспечение экологической безопасности в Мурманской области» на 2009-2010 годы
19	52	Нижегородская область	Концепция областной целевой программы «Экологическая безопасность Нижегородской области на 2011-2015 годы»
20	53	Новгородская область	Областная программа "Охрана окружающей среды и экологическая безопасность области на 2008 - 2010 годы"
21	61	Ростовская область	Областная целевая программа в области охраны окружающей среды и рационального природопользования на 2007-2010 годы
22	64	Саратовская область	Концепция экологической безопасности Саратовской области на 2010-2020 годы
23	66	Свердловская область	Концепция экологической безопасности Свердловской области на период до 2020 года
24	68	Тамбовская область	Долгосрочная целевая программа «Экология и природные ресурсы Тамбовской области на 2009-2012 годы»
25	69	Тверская область	Ведомственная целевая программа "Управление природными ресурсами и охраны окружающей среды Тверской области на 2008 - 2011 годы"
26	71	Тульская область	Областная целевая программа «Экология и природные ресурсы Тульской области на 2007 – 2011 годы»
27	73	Ульяновская область	Областная целевая программа "Охрана

			окружающей среды Ульяновской области на 2007-2010 годы"
28	74	Челябинская область	Концепция основных направлений охраны окружающей среды в Челябинской области на 2007-2015 годы
29	77	г. Москва	Экологическая доктрина города Москвы
30	78	г. Санкт-Петербург	Концепция формирования экологической культуры населения Санкт-Петербурга
31	86	Ханты-Мансийский а. о.	Закон Ханты-Мансийского автономного округа — Югры о целевой программе Ханты-Мансийского автономного округа — Югры "Оздоровление экологической обстановки в Ханты-Мансийском автономном округе — Югре в 2005 - 2010 годах"

Данные в таблице приведены по состоянию на 1 января 2010 г. В регионах, не попавших в таблицу, документов по экологической политике в сети Интернет найдено не было.

Из 31 экологического документа были выбраны задачи и основные направления экологической политики. Каждой задаче (направлению) в зависимости от места, занимаемого в документе, присвоен уровень. Например, главная задача/направление – 1 уровень, подзадача/детализация направления — 2 уровень и т. д. В рассмотренных документах было выделено максимум 3 уровня.

Для наглядности рассмотрим часть задач из Концепции экологической безопасности республика Саха-Якутия (см. Таблицу 2).

В рассматриваемом примере направление 5.2 имеет первый уровень, задачи в рамках направления — второй.

Таблица 2. Уровень и значимость задач

Задача	Уровень	Значимость задачи
...		
5.2. Предотвращение и снижение экологических последствий чрезвычайных ситуаций	1	$\frac{1}{3}$
- своевременное прогнозирование и выявление возможных экологических угроз, включая оценку природных и	2	$\frac{1}{3 \cdot 6}$

техногенных факторов возникновения возможных чрезвычайных ситуаций с негативными экологическими последствиями;		
- обеспечение приоритетного учета интересов населения и качества природной среды при создании новых производств;	2	$\frac{1}{3 \cdot 6}$
- оказание адресной помощи группам населения, проживающим в зонах экологического бедствия, а также особо уязвимым к неблагоприятным экологическим воздействиям (дети, беременные женщины, кормящие матери);	2	$\frac{1}{3 \cdot 6}$
- обучение всех групп населения правилам поведения и способам защиты при чрезвычайных ситуациях с негативными экологическими последствиями;	2	$\frac{1}{3 \cdot 6}$
- поэтапное переселение населения из зон экологического бедствия, техногенных и природных катастроф, не поддающихся реабилитации;	2	$\frac{1}{3 \cdot 6}$
- обеспечение эффективной работы карантинных служб, предотвращение ввоза на территорию республики чужеродных видов, способных нанести вред естественным сообществам и сельскохозяйственным культурам, вредителей, переносчиков и возбудителей заболеваний.	2	$\frac{1}{3 \cdot 6}$
5.3. Предотвращение загрязнения природной среды	1	$\frac{1}{3}$
...		

Для каждой задачи был рассчитан ее вес (значимость) в документе. Значимость задачи n в общем перечне задач объекта равна:

$$M(n) = \frac{1}{k}, \quad (1)$$

где k зависит от уровня задачи.

$M(n)$ рассчитывается как доля данного пункта среди остальных задач. Например, у объекта 5 задач, 2я задача разбивается еще на 3 подзадачи, тогда $M(1), M(2), M(3), M(4), M(5) = 1/5$, а $M(2.1), M(2.2), M(2.3) = 1/15$.

Вернемся к примеру. В концепции выделено 3 направления (в таблице 2 показано одно — 5.2). Направления включают в себя несколько задач, направление 5.2

— 6 задач. Тогда направления концепции Республики Саха имеют значимость $1/3$, а задачи в рамках направления 5.2 значимость $1/3*6$, т. е. $1/18$.

По 31 концепции было выделено 767 задач. Максимальный вес задачи — 0,33, минимальный — 0,0021.

Таблица 3. Сводная таблица

№ п/п	Задача	№ объекта	Вес задачи	Уровень
587	Восстановление и техническое перевооружение государственной службы наблюдения за состоянием окружающей природной среды.	14	0,0185185	2
588	Совершенствование технологий прогнозирования возникновения и развития опасных природных явлений.	14	0,0185185	2
589	Развитие системы обеспечения органов государственной власти, отраслей экономики, Вооруженных Сил и населения оперативной аналитической и режимно-справочной информацией о фактических и ожидаемых изменениях погодно-климатических условий, других гидрометеорологических процессах, а также о состоянии загрязнения окружающей природной среды.	14	0,0185185	2
590	Обеспечение картографической информацией земельно-имущественного и налогового комплексов республики на всех уровнях власти.	14	0,0185185	2
591	Выполнение планового обновления картографической продукции.	14	0,0185185	2
592	Обеспечение выполнения Федерального закона "О наименованиях географических объектов".	14	0,0185185	2
593	Снижение негативного воздействия хозяйственной и иной деятельности на атмосферный воздух и водные объекты.	15	0,2000000	1
594	Снижение негативного воздействия на	15	0,2000000	1

	окружающую среду отходов производства и потребления, включая использование отходов в качестве дополнительных источников сырья.			
595	Минимизация экологических рисков, связанных с возможным возникновением чрезвычайных ситуаций при эксплуатации потенциально аварийно опасных гидротехнических сооружений.	15	0,2000000	1
596	Сохранение и восстановление природной среды.	15	0,2000000	1

2. Группировка задач

Рассматриваемый метод группировки основывается на использовании словаря — тезауруса.

Тезаурус — плоский справочник ключевых словосочетаний, наполнение которого ведется итерационно в рамках алгоритма, представленного ниже. В начале выполнения алгоритма тезаурус пустой.

Таблица 4. Пример тезауруса

№	Словосочетание
1	Вторичные ресурсы
2	Выбросы в атмосферу
3	Экологизация экономики

Исходными данными для выполнения алгоритма являются данные, представленные в таблице 3.

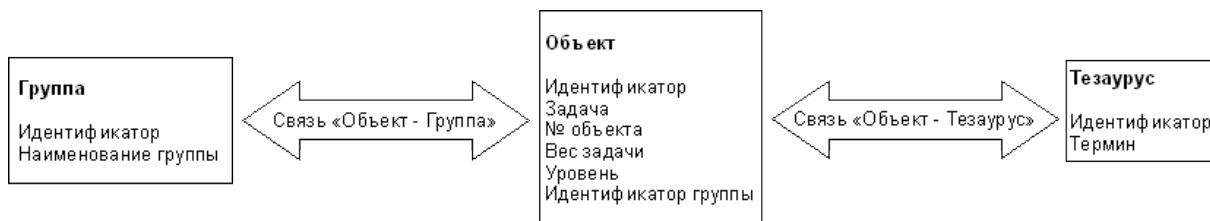


Рис.1. Схема ...

Алгоритм предполагает использование трех таблиц, увязанных между собой связями «Объект — Группа» и «Объект - Тезаурус». В каждой таблице имеется

идентификатор - уникальный номер, однозначно характеризующий запись в таблице

Таблица «Объект» состоит из полей таблицы 3. Добавляется поле «Идентификатор группы».

Таблица «Группа» предназначена для хранения идентификаторов и наименований групп, образующихся в ходе выполнения алгоритма.

Таблица-словарь «Тезаурус» содержит поля «Идентификатор» и «Термин».

Связь «Объект-Группа» считается существующей для объекта, если в таблице «Объект» идентификатор группы равен значению идентификатора из таблицы «Группа».

Связь «Объект-Тезаурус» характеризуется отношением «один ко многим», т. е. одной записи из таблицы «Объект» может соответствовать несколько терминов тезауруса. Эту связь можно представить таблицей:

Таблица 5 . Связь «Объект - Тезаурус»

Идентификатор объекта	Идентификатор тезауруса
3	15
3	107
4	8
4	46
4	33

Группировка проводится по следующему алгоритму:

1. Текст каждой задачи разбивается на ключевые словосочетания, т. е. текстовой формулировке задачи ставится в соответствие набор словосочетаний (их идентификаторов), например 1, 4, 32. Таким образом, текст задачи «Обучение всех групп населения правилам поведения и способам защиты при чрезвычайных ситуациях с негативными экологическими последствиями» можно представить как набор терминов из словаря: Обучение, Население, Чрезвычайные ситуации, Негативные экологические последствия.

Если словосочетание уже имеется в словаре, то в таблице 5 указывается только его идентификатор, если же нет — то оно добавляется в словарь и ему присваивается уникальный номер, который также заносится в таблицу 5.

2. Нахождение наиболее употребимых словосочетаний и группировка задач, которые их содержат. Для этого сначала подсчитывается количество использований

каждого термина из тезауруса (термин не может использоваться в задаче более одного раза).

Таблица 5. Использование терминов тезауруса (по убыванию)

Число использований термина №25	30
Число использований термина №13	23
Число использований термина №105	10
Число использований термина №90	8
...	

Затем выбирается термин, который встречается в текстах задач наибольшее число раз. В нашем случае это термин №25 (см. Таблицу 5), который встречается в 30 задачах. Эти задачи образуют группу — делается запись в таблице «Группа» (см. рис. 1.), задачам присваивается идентификатор группы, например «1». Таким образом заполняется связь «Объект — Группа». Затем удаляются все связи «Объект-Тезаурус», где объект уже попал в группу. Тем самым мы добиваемся того, что каждый объект (задача) попадет только в одну группу. После чего пункт 2 повторяется до тех пор пока не останутся только термины с числом использования в задачах, равным 1.

3. Оптимизация словаря. Если при анализе большого количества задач, появляется надобность переформулировать словосочетание, то поднимаются все тексты задач, содержащие данное словосочетание и, если новая формулировка приемлема в данном случае, ссылка на идентификатор словосочетания сохраняется, если нет — удаляется. Словосочетание получает новое название. Среди текстов задач, которым не подходит новая формулировка словосочетания, т.е. приводит к искажению сути, выделяется новое, более близкое словосочетание, которое добавляется в тезаурус. Редактируется связь «Объект-Тезаурус».

В результате группировки задач концепций, представленных в таблице 1, был получен тезаурус, содержащий 231 термин.

В результате выполнения п.2 алгоритма группировки получено 69 групп, наименования которых можно разделить на три типа: объекты, функции, другие (таблица 6).

Таблица 6. Перечень групп

Идентификатор	Наименование группы	Количество задач в группе	Объект / Функция
1	Окружающая среда (природная среда)	69	о
2	ООПТ (ОПТ)	56	о
3	Водные объекты (ресурсы)	50	о
4	Охрана окружающей среды (природоохранная деятельность)	40	ф
5	Население	36	о
6	Лес	34	о
7	Экологическое образование	33	о
8	Мониторинг	27	ф
9	Технологии	25	о
10	Здоровье населения	21	др
11	Экологическое воспитание (просвещение)	19	ф
12	Отходы	18	о
13	Биоразнообразие	15	др
14	Природные ресурсы	14	о
15	Выброс (сброс)	14	др
16	Почва	14	о
17	Животные и птицы (охрана животных и птиц)	14	о
18	Информационное обеспечение (+информирование)	13	ф
19	Гидротехнические сооружения (ГТС)	13	о
20	Экологические последствия (ущерб природной среде)	11	др
21	Минерально-сырьевые ресурсы (база)	11	о
22	Экологическая безопасность	9	
23	Озелененные территории	9	о
24	Экологическая культура	8	о
25	Техника, оборудование (в т.ч. перевооружение)	7	о/ф
26	Экологическое законодательство	7	о
27	Автотранспорт	6	о

28	Пропаганда (+ реклама) экологических знаний	6	ф
29	Транспорт	6	о
30	Строительство	5	ф
31	Экосистема	5	о
32	Полезные ископаемые	5	о
33	Контроль	5	ф
34	Биологическая безопасность	4	др
35	Экономика (+экономические механизмы)	4	др
36	Воздушные объекты (атмосферный воздух)	4	о
37	Хозяйственная деятельность (хозяйствование)	4	о/ф
38	Природные комплексы	4	о
39	Система управления	4	о
40	Методические материалы, пособия, книги (издательская деятельность)	4	ф
41	Профессиональная подготовка (в т.ч. повышение квалификации)	4	ф
42	Вредное воздействие вод	4	ф
43	Утилизация	3	ф
44	Техногенные аварии, катастрофы	3	о
45	Озеленение	3	ф
46	Реабилитация загрязненных территорий	3	ф
47	Сертификация	3	ф
48	Налоговая политика	3	др
49	Радиационная безопасность	3	др
50	Конкурсы	3	др
51	Экология жилища	3	др
52	Природопользование	2	ф
53	Промышленность	2	о
54	Законодательство природоохранное	2	о
55	Вторичные ресурсы	2	о
56	Естественные сообщества	2	о
57	Опасные вещества (особо опасные, токсичные вещества)	2	о

58	Опасные объекты (экологически)	2	о
59	Экологический каркас	2	о
60	Общественные организации	2	о
61	Негативное воздействие	2	др
62	Продукты питания	2	о
63	Инвентаризация	2	ф
64	Структуры бизнеса	2	о
65	Экологические мероприятия	2	др
66	Экологические проблемы	2	о
67	Картографо-геодезическое обеспечение	2	ф
68	Водные биоресурсы	2	о
69	Заповедники и национальные парки	2	о

3. Построение таблицы суммарных весов задач объектов по группам

Следующим шагом обработки задач экологических концепций для проведения кластеризации является построение таблицы суммарных весов задач объектов по группам.

Рассчитаем количество задач объекта, попавших в группу с учетом весов задач.

$$D_{ij} = \sum (M(n)) = \sum \frac{1}{k} \quad (2)$$

D_{ij} – количество задач объекта i , попавших в группу j с учетом весов этих задач.

$M(n)$ – вес задачи n в общем перечне задач объекта.

k – см. формулу (1).

Т. е. , D_{ij} - сумма весов задач объекта i , попавших в группу j .

Таблица 7. Общий вид таблицы суммарных весов задач объектов по группам

№ группы (j)	1	2	3
№ объекта (i)			
1	D_{11}	D_{12}	D_{13}
2	D_{21}	D_{22}	D_{23}
3	D_{31}	D_{32}	D_{33}
4	D_{41}	D_{42}	D_{43}

В результате получаем таблицу:

Таблица 8. Таблица суммарных весов задач

	1	2	3	...	14	15	16	17	18	19	20	...	30	31
1			0,32	...	0,25			0,14	0,06	0,33	0,13	...		0,6
2			0,01	...	0,16		0,09	0,14	0,06	0,33	0,13	...		
3				...	0,14	0,2			0,12			...		
4		0,25		...	0,12	0,2			0,06			...		
5	0,2		0,26	...				0,29	0,06	0,33		...		
6			0,01	...	0,12		0,09					...		
7	0,2			...			0,09				0,13	...	0,33	
8				...			0,18		0,06			...		
...
68				...	0,01							...		
69				...	0,01							...		

4. Построение матрицы расстояний объектов

Строим матрицу расстояний между объектами, используя данные таблицы суммарных весов задач объектов по группам, полученной на предыдущем шаге.

$$S(x, y) = \frac{\sum (D_{xj} + D_{yj})}{2}, \text{ если } (D_{xj} > 0) \wedge (D_{yj} > 0) \quad (3)$$

$S(x, y)$ – расстояние между двумя объектами.

Таблица 9. Общий вид матрицы расстояний

Объекты	1	2	3	4
1	0	$S(1,2)$	$S(1,3)$	$S(1,4)$
2	$S(1,2)$	0	$S(2,3)$	$S(2,4)$
3	$S(1,3)$	$S(2,3)$	0	$S(3,4)$
4	$S(1,4)$	$S(2,4)$	$S(3,4)$	0

Рассмотрим два крайних значения расстояния между объектами:

Рассмотрим на примере расчет расстояния между двумя объектами. Пусть Объект 1 имеет n задач, а Объект 2 — m задач (рис. 2). Учитывая условие из формулы (3), находим группы, куда попадают задачи с обоих объектов. В нашем случае, это группы g_1, g_5, g_7 .

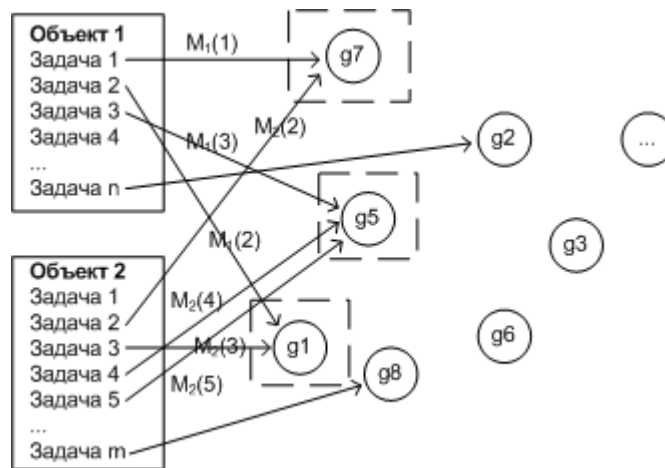


Рис. 2. Пример.

Суммарные значимости задач, входящих в группу g1:

$$D_{11}=M_1(2)$$

$$D_{21}=M_2(3)$$

Суммарные значимости задач, входящих в группу g5:

$$D_{15}=M_1(3)$$

$$D_{25}=M_2(4)+M_2(5)$$

Суммарные значимости задач, входящих в группу g7:

$$D_{17}=M_1(1)$$

$$D_{27}=M_2(2)$$

Рассчитаем расстояние между объектами 1 и 2.

$$S(1,2)=\frac{\sum (D_{xj} + D_{yj})}{2} = \frac{\sum (D_{1j} + D_{2j})}{2} = \frac{(D_{11} + D_{21}) + (D_{15} + D_{25}) + (D_{17} + D_{27})}{2}$$

Расчет матрицы расстояний производится по следующему алгоритму:

В алгоритме используются обозначения:

N – количество кластеров. Перед началом выполнения алгоритма N равно количеству объектов, т. к. каждый объект представляет собой монокластер.

D – пороговое значение диаметра кластера; величина динамическая (см. рис. 3)

step (шаг) — величина в совокупности с номером итерации, определяющая значение d.

Задается исследователем с учетом значений матрицы расстояний.

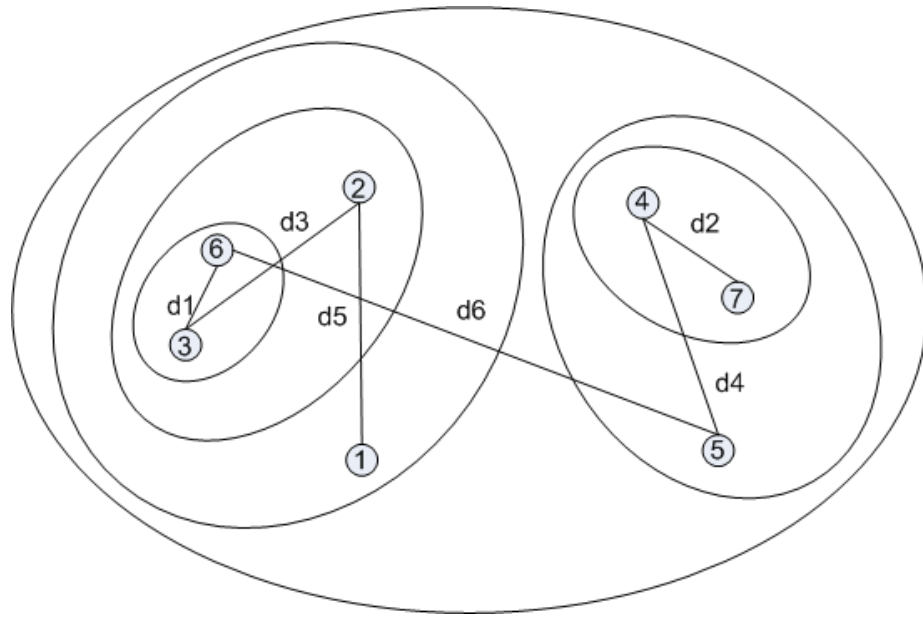


Рис. 3. Кластеризация объектов.

D:=0

start: for D:=D+step do

 for i:= 1 to N

 for (j:= 1...(i-1)) and ((i+1)...N)

 begin

 Рассчитать диаметр кластера i+j: $d(i+j)$ - максимальное из
 расстояний от каждой точки кластера j до
 каждой точки кластера i

 if ($d(i+j) \leq D$) then

 begin

 присоединить кластер j к I

 N:=N-1

 end

 else

 goto start

 end

Результат выполнения алгоритма представлен в таблице 10.

Таблица 10. Исходные данные для построения кластерного дерева

Диаметр d	№ объекта	№ объекта	Название нового
-----------	-----------	-----------	-----------------

			объекта
0,5474568	4	10	A0
0,7850876	8	12	A1
0,8143480	6	A0	A2
0,8192666	14	27	A3
0,9624544	A3	26	A4
1,0241018	11	13	A5
1,1408129	A2	A1	A6
1,2650602	3	19	A7
1,2877678	7	A5	A8
1,3749998	28	22	A9
1,3930350	A4	A6	A10
1,4093961	A7	17	A11
1,5789472	31	25	A12
1,7762116	23	A8	A13
1,8827875	18	A10	A14
2,0581931	2	A13	A15
2,0952381	16	20	A16
2,1000000	29	21	A17
2,1428570	5	A12	A18
2,1538462	9	A11	A19
2,3807274	A9	A14	A20
2,9166666	24	1	A21
4,0643234	A20	15	A22
4,2000000	A21	30	A23
4,2000000	A18	A17	A24
4,2000000	A24	A19	A25
6,8750002	A23	A16	A26

II. Результаты и их обоснование

На основе табличных данных строим древовидную кластерную диаграмму. По оси X откладываются идентификационные номера объектов, по оси Y – диаметры кластеров — рис. 4.

На основе кластерного дерева можно выделить приведенные в таблице 11 группы концепций. В скобках во второй колонке указано количество объектов, у

которых имеются задачи, входящие в данную группу. Через дробь указано количество объектов в группе.

Таблица 11. Группы концепций

№ п\п	Общие положения документов	Регионы в группе
1.	Водные объекты (12/12), Окружающая среда (11/12), ООПТ (9/12), Охрана окружающей среды (9/12)	Республика Татарстан, Свердловская, Архангельская, Нижегородская области, г. Москва, Республика Карелия, Тамбовская, Ростовская, Воронежская, Тверская, Курганская области, Чувашская республика.
2.	Охрана окружающей среды (5/5), Окружающая среда (4/5), ООПТ (4/5), Природные ресурсы (4/5), Технологии (4/5)	Челябинская область, г. Санкт-Петербург, Астраханская, Курская области, р-ка Дагестан.
3.	Окружающая среда (9/9), Население (5/9), ООПТ (3/9)	Республика Саха (Якутия), Камчатская, Вологодская, Саратовская, Тульская, Костромская области, Ханты-Мансийский а.о., Новгородская область, Республика Тува.
4.	Экологическое образование (5/5), ООПТ (3/5)	Ставропольский край, Кировская, Мурманская области, Республика Бурятия, Ульяновская область.

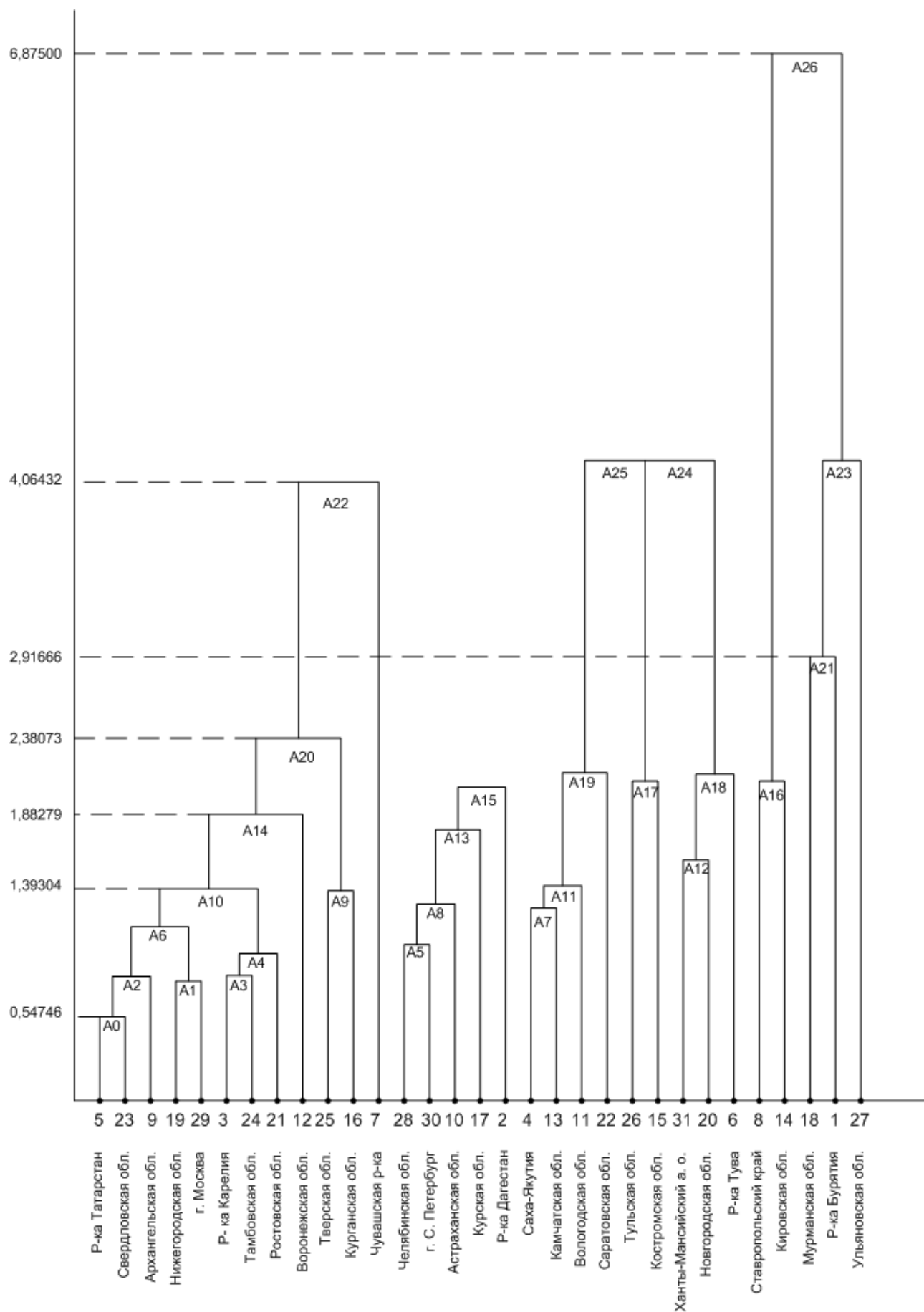


Рис. 4. Кластерное дерево.

Введем понятие основной группы.

Основная группа — группа, в которую попадает хотя бы одна задача от каждого объекта (концепции).

На основе введенного понятия проследим изменение количества основных групп на разных уровнях d.

Для группы №1 (см. таблицу 11) имеем:

Таблица 12. Изменение количества основных групп. Группа №1.

Наименование группы	A2	A6	A10	A14	A20	A22
Водные объекты (ресурсы)	+	+	+	+	+	+
Здоровье населения	+	+				
Контроль	+	+				
Мониторинг	+	+				
Население	+	+				
Окружающая среда (природная среда)	+	+	+	+	+	
ООПТ (ОПТ)	+	+	+	+		
Почва	+					
Природные ресурсы	+					
Охрана окружающей среды (природоохранная деятельность)	+					
Технологии	+					
Экологическая безопасность	+					
Экологические последствия (ущерб природной среде)	+					

Таблица 13. Изменение количества основных групп. Группа №2.

Наименование группы	A5	A8	A13	A15
Население	+			
Окружающая среда (природная среда)	+	+	+	
ООПТ (ОПТ)	+	+	+	
Охрана окружающей среды (природоохранная деятельность)	+	+	+	+
Технологии	+	+		
Экологическое воспитание (просвещение)	+			
Экологическое образование	+	+		

Таблица 14. Изменение количества основных групп. Группа №3.

Наименование группы	A7	A11	A25
Окружающая среда (природная среда)	+	+	+
ООПТ (ОПТ)	+	+	
Население	+	+	

Таблица 15. Изменение количества основных групп. Группа №4.

Наименование группы	A16	A26
Вредное воздействие вод	+	
Минерально-сырьевые ресурсы (база)	+	
ООПТ (ОПТ)	+	
Экологическое образование	+	+

III. Итоги

В статье был рассмотрен метод классификации экологических концепций регионов, который позволяет разделить экологические документы регионов на группы. В выделенных группах, как показал анализ результатов, прослеживается основное направление экологической политики. При решении вопроса поиска регионов с близкими экологическими задачами метод позволяет позиционировать исследуемый регион, для которого составляется экологическая концепция.

Построенная кластерная диаграмма наглядно отображает динамику образования кластеров объектов.

Рассмотренный метод включает следующие этапы:

1. Выделение из текста существенных признаков - задач.
2. Группировка задач с использованием тезауруса.
3. Построение таблицы суммарных весов задач объектов по группам.
4. Построение матрицы расстояний объектов.
5. Построение кластерного дерева и выделение групп концепций.

Таким образом, существующие сегодня концепции экологической политики регионов могут быть разделены на 4 группы по следующим направлениям:

1. Водные объекты.
2. Охрана окружающей среды.
3. Окружающая среда.
4. Экологическое образование.

Описанный метод может применяться для анализа других объектов, имеющих стратегические документы. Авторы проанализировали близость ВУЗов г. Самары на основе стратегий по информатизации, в результате были получены актуальные данные, позволившие определить группы, тяготеющие к тем или иным задачам информатизации в ВУЗах и эффективно организовать обмен опытом.

Таким образом, предложенный метод позволяет обобщить опыт развития определенной группы объектов и выбрать наиболее подходящий для конкретного объекта набор прототипов из имеющегося множества.