Министерство образования и науки Российской Федерации

Федеральное государственное бюджетное образовательное учреждение

высшего профессионального образования

УЛЬЯНОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

УТВЕРЖДАЮ

Заведующий кафедрой

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(подпись)

\_Крашенинников В.Р.\_

( ФИО )

“\_\_\_” \_\_июня\_ 2014 г.

**ДИПЛОМНАЯ РАБОТА**

Разработка программы кластеризации ИТ-предприятий

(наименование работы)

Автор работы \_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_Сибирев И.В., ПМд-51\_

(подпись) (ФИО, группа)

Руководитель работы \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_Афанасьева Т.В.\_

(подпись) (ФИО)

Рецензент \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(подпись) (ФИО)

Ульяновск

2014 г.

Министерство образования и науки Российской Федерации

Федеральное государственное бюджетное образовательное учреждение

высшего профессионального образования

УЛЬЯНОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

УТВЕРЖДАЮ

Заведующий кафедрой

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(подпись)

\_Крашенинников В.Р.\_

( ФИО )

“\_\_\_” июня 2014 г.

**ЗАДАНИЕ НА ДИПЛОМНУЮ РАБОТУ**

студенту группы ПМд-51 экономико-математического факультета

\_\_\_\_\_\_\_\_\_\_\_\_\_Сибиреву И.В.\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(Фамилия, И.О.)

1.Тема: Разработка программы кластеризации ИТ-предприятий

утверждена приказом по УлГТУ №от « » февраля\_ 2014 г.

2. Срок представления работы к защите “\_14\_” июня 2014 г.

3. Исходная информация к работе

Методы кластерного анализа, теория нечетких множеств, экономические модели сегментации рынка; методические, аналитические, научно-практические материалы, результаты статистических и социологических исследований ИТ-кластера.

4. Содержание работы

Описание модели кластеризации ИТ- предприятий, анализ и выбор критериев и параметров, среди которых имеются лингвистические данные, используемых при кластеризации ИТ- предприятий, разработка алгоритма кластеризации, разработка программы кластеризации ИТ-предприятий с использованием иерархического центроидного и неиерархического FCM-метода нечеткой кластеризации, применение программы при обработке данных ИТ-компаний Ульяновской области

5. Графические материалы

Презентация, раздаточный материал \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

6. Приложения

Текст программы\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Руководитель работы \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_Афанасьева Т.В.\_\_

(подпись) (ФИО)

“\_\_\_” \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 2014 г.

Консультант \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(подпись) (ФИО)

Задание принял к исполнению “\_\_\_” \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 2014 г. \_\_\_\_\_\_\_\_\_\_\_\_\_\_

(подпись студента)

УТВЕРЖДАЮ

Руководитель работы

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(подпись)

Афанасьева Т.В.\_

(ФИО)

«\_\_\_»февраля 2014 г.

#### **КАЛЕНДАРНЫЙ ПЛАН**

выполнения дипломной работы

по теме Разработка программы кластеризации ИТ-предприятий

студентом группы ПМд - 51 Сибирева Ивана Валерьевича

(ФИО)

|  |  |
| --- | --- |
| **Вид работы** | **сроки** |
| Изучение литературы, работа с Internet-источниками, обзор литературы по теме работы, изучение возможностей программ подобного типа | 17.10 |
| Создание и тестирование ранней версии программы «Кластеризация сложных объектов». Подготовка презентации | 13.11 |
| Апробация. Участие в конкурсе инновационных проектов по программе Фонда содействия развитию малых форм предприятий в научно-технической сфере «Участник молодежного научно-инновационного конкурса 2013» («УМНИК») | 19.11 |
| Оформление курсовой работы «Программа кластеризация сложных объектов». | 20.12 |
| Описание модели кластеризации ИТ- предприятий. | 30.01 |
| Подготовка статьи | 10.02 |
| Изучение методических, аналитических, научно-практических материалов, результатов статистических и социологических исследований ИТ-кластера, полученных коллективом ученых УлГТУ.  Анализ объектов ИТ-рынка. Анализ и выбор критериев и параметров, среди которых имеются лингвистические данные, используемых при кластеризации ИТ- предприятий. | 12.03 |
| Разработка алгоритма кластеризации ИТ-предприятий иерархическим центроидным и неиерархическим FCM-методом нечеткой кластеризации | 12.04 |
| Разработка программы кластеризации ИТ-предприятий, применение программы при обработке данных ИТ-компаний Ульяновской области | 30.04 |
| Обработка и анализ данных. Унификация лингвистических данных из анкет, представленных руководителями ИТ-предприятий Ульяновской области | 15.05 |
| Тестирование и отладка программы. Экспериментальная часть:кластеризация ИТ-предприятий рынка Ульяновской области;анализ текущего уровня развития ИТ-предприятий в регионе. | 15.05 |
| Анализ результатов. Подготовка выводов и рекомендаций | 17.05 |
| Оформление текста работы | 23.05 |
| Оформление графических материалов | 30.05 |
| Рассмотрение работы на заседании кафедры | .06 |
| Редактирование и доработка дипломной работы | 10.06 |
| Рецензирование | 18.06 |
| Защита работы | 20.06 |

Студент \_\_Сибирев И.В.\_\_

(ФИО)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(подпись)

«\_\_\_ » февраля 2014 г.

**Реферат**

Дипломная работа посвящена разработке программы кластеризации ИТ-предприятий. Целью кластерного анализа является разбиение множества исследуемых объектов и признаков на однородные группы или кластеры, т.е. решается задача классификации данных и выявления соответствующей структуры в ней. Актуально сочетание методов кластерного анализа и теории нечетких множеств.

Актуальным является исследование текущего уровня ИТ-сектора. Этой тематикой занимается коллектив ученых УлГТУ: Н.Г. Ярушкина, Т.В. Афанасьева, О.В. Шиняева, К.В. Святов, Л.М. Валкин, Д.А. Ефремов, К.Г. Калимуллин и др. Результаты исследования представлены в монографии Н.Г. Ярушкиной, Т.В. Афанасьевой и др. «Исследование ИТ-кластера Ульяновской области». На основе предложенной в монографии методики выбора критериев и параметров, характеризующих ИТ-сектор, в работе производится кластеризация ИТ-предприятий Ульяновской области.

**Цель работы** – создать программный инструмент для кластеризации ИТ-предприятий.

Нами разработан программный продукт «Кластеризация ИТ-предприятий». Программный продукт позволяет группировать данные иерархическим центроидным методом кластерного анализа, неиерархическим FCM-методом, который предполагает возможность работы с нечеткими данными. Выполнена программная реализация методов создания, сохранения, загрузки, обработки кластеров (данных).

Программный продукт позволяет: визуализировать данные на экране в виде дерева, позволяющего получить результат кластеризации в виде иерархической структуры, по которой можно судить о ходе выполнения кластеризации, выводить на экран результаты анализа данных в виде таблиц, графиков и диаграмм.

Используется программный код, который может выполняться на различном HardWare, без его предварительного переноса на «родной язык» устройства. Нами был выбран C#. Код на C# будет работать везде, где установлен «C# .NET». Сохранение и загрузка данных происходит в формате \*.SOF (Serializable Object Files).

Предусмотрена возможность использования кластерного анализа в сочетании с другими методами многомерного анализа, возможность использования различных метрик; возможность включения программного продукта в состав более сложной автоматической или автоматизированной системы для экономического анализа.

Подобных программных продуктов для сегментации рынка IT нами не найдено.

**Практическая значимость.** Программный продукт предназначен для кластерного анализа данных IT-предприятий, для обработки экономических данных и их анализа, группировки, распознавания и представления в удобном для пользователя виде.

Программное обеспечение предполагается использовать для кластеризации ИТ-предприятий по следующим группам критериев: область деятельности компании; финансово-экономические показатели; рынок сбыта; кадровые показатели; внутренние процессы компании; инновационная деятельность; социальные программы и образовательная деятельность. Программное обеспечение адаптировано для экономической терминологии «Кластеризация IT-предприятий».

Потребителями данного программного продукта могут выступить экономисты, социологи, фирмы, лица, занимающиеся производством и продажей в сфере IT, которым полезно сегментирование рынка товаров и услуг, структурирование рынка, основанное на неоднородности и нечеткости показателей и критериев. Такой анализ необходим для дифференцированного маркетинга.

**Апробация.** Программный продукт был представлен на конкурсе инновационных проектов по программе Фонда содействия развитию малых форм предприятий в научно-технической сфере «Участник молодежного научно-инновационного конкурса 2013» («УМНИК»).

Работа представлена на XXIX международной научно-практической конференции «Инновации в науке» (Новосибирск, 2014).

Опубликована статья Афанасьевой Т.В., Сибирева И.В. «Программа «Сегментации и кластеризация рынка IT» // Инновации в науке. № 1 (26). –Новосибирск: Изд. «СибАК», 2014. С. 46-53.

Работа состоит из введения, четырех глав, заключения, списка литературы и приложения.

Первая глава посвящена истории возникновения и развития кластерного анализа и теории нечетких множеств, здесь рассмотрены основные понятия, приведен обзор методов кластерного анализа.

Во второй главе описана методика кластеризации ИТ-предприятий: выбор критериев и параметров, используемых при кластеризации ИТ-предприятий; анализ и предварительная обработка данных анкет руководителей ИТ-предприятий г. Ульяновска; методы и алгоритмы кластеризации, применяемые при написании программы «Кластеризация ИТ-предприятий».

В третьей главе произведено описание программного продукта «Кластеризация ИТ-предприятий», его функций, руководство пользователю, руководство программисту.

В четвертой главе произведен анализ результатов кластеризации ИТ-предприятий Ульяновской области. При вычислительном эксперименте используются результаты анкетировании руководителей ИТ-предприятий в 2013 г.

**Оглавление**

[Введение 9](#_Toc390634819)

[Глава 1. Теоретико-методологические основы изучения многомерных данных сложных систем 15](#_Toc390634820)

[§ 1.1. Возникновение и развитие кластерного анализа и теории нечетких множеств 15](#_Toc390634821)

[§ 1.3. Введение в теорию нечетких множеств 26](#_Toc390634822)

[§ 1.4. Нечеткие системы кластеризации 30](#_Toc390634823)

[§ 1.5. Обзор программных продуктов для кластеризации многомерных данных. 34](#_Toc390634824)

[Выводы 37](#_Toc390634825)

[Глава 2. Методика кластеризации ИТ-предприятий 39](#_Toc390634826)

[2.1 Выбор критериев и параметров, используемых при кластеризации ИТ-предприятий 39](#_Toc390634827)

[§ 2.2. Анализ и предварительная обработка данных анкет руководителей ИТ-предприятий г. Ульяновска 47](#_Toc390634828)

[§ 2.3. Методы кластеризации, применяемые при написании программы «Кластеризация ИТ-предприятий» 56](#_Toc390634829)

[Выводы. 61](#_Toc390634830)

[Глава 3. Описание программного продукта «Кластеризация ИТ-предприятий» 62](#_Toc390634831)

[§ 3.1. Описание функций программного продукта «Кластеризации ИТ-предприятий» 62](#_Toc390634832)

[§ 3.2. Руководство пользователю 64](#_Toc390634833)

[§ 3.3. Руководство программисту 73](#_Toc390634834)

[Выводы 76](#_Toc390634835)

[Глава 4. Анализ результатов кластеризации ИТ-предприятий Ульяновской области 77](#_Toc390634836)

[§ 4.1. Результаты кластеризации по критерию «Область деятельности компании» 77](#_Toc390634837)

[§ 4.2. Кластеризация по финансово-экономическим показателям 80](#_Toc390634838)

[§ 4.3 Кластеризация по критерию «Рынок сбыта» 85](#_Toc390634839)

[§ 4.4. Кадровые показатели. 88](#_Toc390634840)

[§ 4.5. Внутренние процессы компании. Инновационная деятельность. Социальные программы и образовательная деятельность. 90](#_Toc390634841)

[§ 4.6. Кластеризация по всем группам параметров. 95](#_Toc390634842)

[§ 4.7. Кластеризация ИТ-предприятий Ульяновской области FCM-методом. Проверка надежности и достоверности результатов кластеризации. 1](#_Toc390634843)

[Выводы. 2](#_Toc390634844)

[Заключение. 3](#_Toc390634845)

[Литература 4](#_Toc390634846)

[Приложение 7](#_Toc390634847)

[Приложение 1 1](#_Toc390634848)

# Введение

Одним из приоритетных направлений модернизации экономики Российской Федерации является реализация кластерного подхода. С 2011 года в России идет активная кластерная политика под эгидой Министерства экономического развития РФ, принято решение поддержать часть кластеров через предоставление субсидий из федерального бюджета.

Согласно определению, приведенному в «Методических рекомендациях по реализации кластерной политики» Минэкономразвития России, кластер – это «объединение предприятий, поставщиков оборудования, комплектующих, специализированных производственных и сервисных услуг, научно-исследовательских и образовательных организаций, связанных отношениями территориальной близости и функциональной зависимости в сфере производства и реализации товаров и услуг».

Кластерные хозяйственные системы отличаются высокой конкурентоспособностью и стабильностью. В Ульяновской области развиваются ядерный, автомобильный, авиационный, пищевой, стекольный кластеры, кластер производства строительных материалов [22].

Одним из приоритетных направлений развития экономики является развитие информационных технологий, поэтому актуально формирование ИТ-кластера Ульяновской области, где происходит разработка новых информационных технологий и программного обеспечения, их реализация и внедрение в производство. В связи с этим актуальным является исследование текущего уровня ИТ-сектора, применение методов кластерного анализа, сравнение теоретического анализа данных с намечающимися кластерами ИТ-предприятий.

Этой тематикой занимается коллектив ученых УлГТУ: Н.Г. Ярушкина, Т.В. Афанасьева, О.В. Шиняева, К.В. Святов, Л.М. Валкин, Д.А. Ефремов, К.Г. Калимуллин и др. Результаты исследования представлены в монографии Н.Г. Ярушкиной, Т.В. Афанасьевой и др. «Исследование ИТ-кластера Ульяновской области» [22].

Социально-экономические системы в их динамическом развитии характеризуются многомерными данными. При анализе информации о сложных системах необходимы обработка и анализ огромных объёмов разнородных данных, важно упорядочить их по иерархическим уровням и произвести группировку, разбиение множества исследуемых объектов и признаков на однородные группы или кластеры. Это позволяет сделать кластерный анализ, впервые применённый в социологии Трионом [25] в 1939г.

Встаёт проблема автоматизации процесса обработки и анализа информации, необходимости создания простого и доступного инструмента (программного пакета) для группировки и распознавания данных, произвольно структурированных, и типов данных. Остро возникает проблема представления данных в виде удобном и понятном для пользователя. Необходим программный код, который бы мог выполняться на различном HardWare, без его предварительного переноса на «родной язык» устройства. Актуально использование методов кластерного анализа для численных и нечетких данных.

На сегодняшний день актуальным является развитие методов кластерного анализа (см. [9],[10], [11], [12], [15] и др.), среди которых метод полных связей, метод максимального локального расстояния, метод Ворда, центроидный метод и др. Актуально их применение к решению экономических задач (см. [6], [17], [18], [20] и др.), в частности, сегментации рынка (модель Уинда и Кардозы и др.), их программная реализация.

На сегодняшний день актуально сочетание кластерного анализа с теорией нечетких множеств (см. [2], [3], [5], [7], [13], [21], [23] и др.), основы которой заложены в 60-е годы американским математиком Латфи Заде [5]. В настоящее время Японией и США активно разрабатываются электронные системы с нечеткими управляющими алгоритмами.

Существует программное обеспечение для кластерного анализа в узких предметных областях, например, «ClusterDelta» позволяет производить анализ данных на бирже, существуют программы для прогноза конъюнктуры рынка и др.

Существуют пакеты программ, реализующие набор методов кластерного анализа данных («STATISTICA Multivariate Exploratory Techniques», «MatLab» и др.) В программном пакете «STATISTICA Multivariate Exploratory Techniques» реализованы методы k-средних, иерархической кластеризации и двухвходового объединения. Используются различные метрики расстояний. Этот пакет не ориентирован на решение задач кластеризации в условиях неоднородных показателей, на обработку нечетких данных. В «MatLab» имеются встроенные функции, позволяющие реализовать алгоритмы четкой и нечеткой кластеризации. Эти пакеты рассчитаны на использование профессиональными математиками и в этом их ограничение.

**Цель работы** – создать программный инструмент для кластеризации ИТ-предприятий.

**Задачи работы:**

– изучение методов кластерного анализа, элементов теории нечетких множеств;

– изучение возможностей, особенностей программной реализации в программных продуктах для кластерного анализа;

– программная реализация методов создания, сохранения, загрузки, обработки кластеров (данных);

– изучение методических, аналитических материалов, результатов статистических и социологических исследований ИТ-кластера;

– анализ объектов ИТ-рынка, анализ и выбор критериев и параметров, используемых при кластеризации ИТ-предприятий;

– разработка алгоритма кластеризации ИТ-предприятий; разработка программы кластеризации ИТ-предприятий;

– экспериментальная часть: применение программы при обработке данных ИТ-компаний Ульяновской области, кластеризация ИТ-предприятий Ульяновской области; анализ текущего уровня развития ИТ-предприятий в регионе.

Нами разработан программный продукт «Кластеризация ИТ-предприятий». Программный продукт позволяет группировать данные иерархическим центроидным методом кластерного анализа, неиерархическим FCM-методом, который предполагает возможность работы с нечеткими данными. Выполнена программная реализация методов создания, сохранения, загрузки, обработки кластеров (данных).

Программный продукт позволяет: визуализировать данные на экране в виде дерева, позволяющего получить результат кластеризации в виде иерархической структуры, по которой можно судить о ходе выполнения кластеризации, выводить на экран результаты анализа данных в виде таблиц, графиков и диаграмм.

Используется программный код, который может выполняться на различном HardWare, без его предварительного переноса на «родной язык» устройства. Нами был выбран C#. Код на C# будет работать везде, где установлен «C# .NET». Сохранение и загрузка данных происходит в формате \*.SOF (Serializable Object Files).

Предусмотрена возможность использования кластерного анализа в сочетании с другими методами многомерного анализа, возможность использования различных метрик; возможность включения программного продукта в состав более сложной автоматической или автоматизированной системы для экономического анализа.

Программа может быть интегрирована в состав более сложной автоматической или автоматизированной вычислительной системы (автоматической системы управления, прогнозирования и т.д.), размеры которой могут быть как в пределах одной машины, в пределах локальной, глобальной сети, с оповещением через уже существующие средства связи (sms, Email, …).

**Новизна работы** в том, что:

– подобных программных продуктов для сегментации рынка IT нами не найдено;

– используется авторский подход программной реализации методов кластерного анализа;

– сохранение и загрузка данных происходит в формате \*.SOF (Serializable Object Files), пользователь имеет возможность редактировать данные вне программы;

– возможность использования кластерного анализа в сочетании с другими методами многомерного анализа.

– программное обеспечение ориентировано на обработку числовых и нечетких данных.

**Практическая значимость.** Программный продукт предназначен для кластерного анализа данных IT-предприятий, для обработки экономических данных и их анализа, группировки, распознавания и представления в удобном для пользователя виде.

Программное обеспечение предполагается использовать для сегментации ИТ-предприятий Ульяновской области по следующим группам критериев: область деятельности компании; финансово-экономические показатели; рынок сбыта; кадровые показатели; внутренние процессы компании; инновационная деятельность; социальные программы и образовательная деятельность. При вычислительном эксперименте используются результаты анкетировании руководителей ИТ-предприятий в 2013 г. [21].

Программное обеспечение адаптировано для экономической терминологии «Кластеризация IT-предприятий».

Потребителями данного программного продукта могут выступить экономисты, социологи, фирмы, лица, занимающиеся производством и продажей в сфере IT, которым полезно сегментирование рынка товаров и услуг, структурирование рынка, основанное на неоднородности и нечеткости показателей и критериев. Такой анализ необходим для дифференцированного маркетинга.

**Апробация.** Программный продукт был представлен на конкурсе инновационных проектов по программе Фонда содействия развитию малых форм предприятий в научно-технической сфере «Участник молодежного научно-инновационного конкурса 2013» («УМНИК»).

Работа представлена на XXIX международной научно-практической конференции «Инновации в науке» (Новосибирск, 2014).

Опубликована статья Афанасьевой Т.В., Сибирева И.В. «Программа «Сегментации и кластеризация рынка IT» [4].

Работа состоит из введения, четырех глав, заключения, списка литературы и приложения.

Первая глава посвящена истории возникновения и развития кластерного анализа и теории нечетких множеств, здесь рассмотрены основные понятия, приведен обзор методов кластерного анализа.

Во второй главе описана методика кластеризации ИТ-предприятий: выбор критериев и параметров, используемых при кластеризации ИТ-предприятий; анализ и предварительная обработка данных анкет руководителей ИТ-предприятий г. Ульяновска; методы и алгоритмы кластеризации, применяемые при написании программы «Кластеризация ИТ-предприятий».

В третьей главе произведено описание программного продукта «Кластеризация ИТ-предприятий», его функций, руководство пользователю, руководство программисту.

В четвертой главе произведен анализ результатов кластеризации ИТ-предприятий Ульяновской области.

Библиография содержит 47 литературных источников, в том числе 22 – электронных ресурса.

# Глава 1. Теоретико-методологические основы изучения многомерных данных сложных систем

## § 1.1. Возникновение и развитие кластерного анализа и теории нечетких множеств

При анализе и прогнозировании социально-экономических явлений исследователь сталкивается с многомерностью их описания. Это  происходит при решении задачи сегментирования рынка, построении типологии стран по большому числу показателей, прогнозирования конъюнктуры рынка товаров, изучении и прогнозировании экономической депрессии и др.

Методы многомерного анализа – действенный количественный инструмент исследования социально-экономических процессов, описываемых большим  числом характеристик. К ним относятся кластерный анализ, таксономия, распознавание образов, факторный анализ и др. Кластерный анализ наиболее ярко отражает черты многомерного анализа в классификации, факторный анализ – в исследовании связи.

Первое применение кластерный анализ нашел в социологии. Название кластерный анализ происходит от английского слова cluster – гроздь, скопление. Впервые в 1939 был определен предмет кластерного анализа и сделано его описание исследователем Трионом. Главное назначение кластерного анализа – разбиение множества исследуемых объектов и признаков на однородные группы или кластеры. Решается задача классификации данных и выявления соответствующей структуры в ней.

Большое достоинство кластерного анализа в том, что он позволяет производить разбиение объектов не по одному параметру, а по целому набору признаков. Кластерный анализ в отличие от большинства математико-статистических методов не накладывает никаких ограничений на вид рассматриваемых объектов, и позволяет рассматривать множество исходных данных практически произвольной природы. Это имеет большое значение, например, для прогнозирования конъюнктуры, когда показатели имеют разнообразный вид, затрудняющий применение традиционных эконометрических подходов.

Кластерный анализ позволяет рассматривать достаточно большой объем информации и резко сокращать, сжимать большие массивы социально-экономической информации, делать их  компактными и наглядными.

Кластерный анализ можно использовать циклически. В этом случае исследование производится до тех пор, пока не будут достигнуты необходимые результаты. При этом каждый цикл здесь может давать информацию, которая способна сильно изменить направленность и подходы дальнейшего применения кластерного анализа. Этот процесс можно представить системой с обратной связью.

В задачах  социально-экономического прогнозирования перспективно сочетание кластерного анализа  с другими количественными методами (например, с регрессионным анализом).

Кластерный анализ имеет недостатки  и ограничения. В частности, состав  и количество кластеров зависит от  выбираемых критериев разбиения. При сведении исходного массива данных к более компактному виду могут возникать искажения, теряться индивидуальные черты отдельных объектов за счет  замены их характеристиками обобщенных значений параметров кластера. При проведении классификации часто игнорируется возможность отсутствия в рассматриваемой  совокупности каких-либо значений.

Существует достаточно много методов кластерного анализа. Например, метод полных связей, метод максимального локального расстояния, метод Ворда, центроидный метод (методом взвешенных групп) и др.

Алгоритмы кластерного анализа можно  подразделить на иерархические и неиерархические. Иерархические алгоритмы делятся на агломеративные, характеризуемые последовательным объединением  исходных элементов и соответствующим уменьшением числа кластеров, и на дивизимные (делимые), в которых число кластеров возрастает, начиная с одного, в результате чего образуется последовательность  расщепляющих групп.

Алгоритмы кластерного анализа имеют хорошую программную реализацию, которая позволяет решить задачи самой большой размерности.

Сочетание методов кластерного анализа с методами теории нечетких множеств открывает новые возможности для исследователей.

Основы нечеткой логики были заложены в конце 60-х лет в работах известного американского математика Латфи Заде [5]. Исследования такого рода было вызвано возрастающим неудовольствием экспертными системами. «Искусственный интеллект» легко справлялся с задачами управления сложными техническими комплексами, но был беспомощным при простейших высказываниях повседневной жизни.

Для создания действительно интеллектуальных систем, способных адекватно взаимодействовать с человеком, был необходим новый математический аппарат, который переводит неоднозначные жизненные утверждения в язык четких и формальных математических формул. Первым серьезным шагом в этом направлении стала теория нечетких множеств, разработанная Заде. Его работа «Fuzzy Sets», опубликованная в 1965 году в журнале «Information and Control», заложила основы моделирования интеллектуальной деятельности человека и стала начальным толчком к развитию новой математической теории. Он же дал и название для новой области науки –«fuzzy logic» (fuzzy - нечеткий, размытый, мягкий).

Дальнейшие работы профессора Латфи Заде и его последователей заложили фундамент новой теории и создали предпосылки для внедрения методов нечеткого управления в инженерную практику.

Аппарат теории нечетких множеств, продемонстрировав ряд многообещающих возможностей применения – от систем управления летательными аппаратами до прогнозирования итогов выборов, оказался вместе с тем сложным для практического воплощения. Нечеткая логика заняла свое место среди других специальных научных дисциплин – где-то посредине между экспертными системами и нейронными сетями.

Свое второе рождение теория нечеткой логики пережила в начале восьмидесятых годов, когда несколько групп исследователей (в основном, в США и Япони) всерьез занялись созданием электронных систем различного применения, использующих нечеткие управляющие алгоритмы. Теоретические основы для этого были заложены в ранних работах Коско и других ученых.

Третий период начался с конца 80-х годов и до сих пор. Этот период характеризуется бумом практического применения теории нечеткой логики в разных сферах науки и техники. До 90-ого года появилось около 40 патентов, относящихся к нечеткой логике (30 - японских). Сорок восемь японских компаний создают лабораторию LIFE (Laboratory for International Fuzzy Engineering), японское правительство финансирует 5-летнюю программу по нечеткой логике, которая включает 19 разных проектов - от систем оценки глобального загрязнения атмосферы и предвидения землетрясений до АСУ заводских цехов. Результатом выполнения этой программы было появление целого ряда новых массовых микрочипов, базирующихся на нечеткой логике. Сегодня их можно найти в стиральных машинах и видеокамерах, цехах заводов и моторных отсеках автомобилей, в системах управления складскими роботами и боевыми вертолетами.

В США развитие нечеткой логики идет по пути создания систем для большого бизнеса и военных. Нечеткая логика применяется при анализе новых рынков, биржевой игре, оценки политических рейтингов, выборе оптимальной ценовой стратегии и т.п. Появились и коммерческие системы массового применения.

Смещение центра исследований нечетких систем в сторону практических применений привело к постановке целого ряда проблем, в частности: новые архитектуры компьютеров для нечетких вычислений; элементная база нечетких компьютеров контроллеров; инструментальные средства разработки; разработка нечетких систем управления, и т.п.§ 1.2. Многомерный кластерный анализ

**Задача кластерного анализа.** Задача кластерного анализа заключается в том, чтобы на основании данных, содержащихся во множестве , разбить множество объектов на ( – целое) кластеров (подмножеств) , , …, , так, чтобы каждый объект принадлежал одному и только одному подмножеству разбиения и чтобы объекты, принадлежащие одному и тому же кластеру, были сходными, в то время, как объекты, принадлежащие разным кластерам были разнородными (см. [9],[10], [11], [12], [15] и др.).

Решением задачи кластерного анализа являются разбиения, удовлетворяющие некоторому критерию оптимальности. Этот критерий может представлять собой некоторый функционал, который называют целевой функцией. Например, в качестве целевой функции может быть взята внутригрупповая сумма квадратов отклонения:

,

где - представляет собой измерения -го объекта.

Для решения  задачи кластерного анализа необходимо определить понятие сходства и разнородности.

Объекты -ый и -ый попадут в один кластер, когда расстояние между точками и будет достаточно мало и попадут в разные кластеры, когда это расстояние будет достаточно большим. Попадание в один или разные кластеры объектов определяется понятием расстояния между и из , где – -мерное евклидово пространство.

Неотрицательная функция называется функцией расстояния (метрикой), если:

а) для всех и из ;

б) тогда и только тогда, когда ;

в) ;

г) ;

где – любые три вектора из .

Наиболее часто употребляются следующие функции расстояний:

евклидово расстояние        ,

– норма                            ,

супремум - норма              , ,

– норма                             .

Евклидова метрика является наиболее популярной. Метрика наиболее легкая для вычислений. Супремум-норма легко считается и включает в себя процедуру упорядочения, а - норма обобщает три первые функции расстояний.

Пусть n измерений , ,..., представлены в виде матрицы данных размером :

.

Тогда расстояние между парами векторов могут быть представлены в виде симметричной матрицы расстояний:

.

Важным понятием является понятие сходства между объектами и . Неотрицательная вещественная функция  называется мерой сходства, если:

1) для ;

2) ;

3) ;

Пары значений мер сходства можно объединить в матрицу сходства:

.

Величину называют коэффициентом сходства.

Естественной  мерой сходства характеристик объектов во многих задачах является коэффициент корреляции между ними



где , , , соответственно средние и среднеквадратичные отклонения для характеристик и . Мерой различия между характеристиками может служить величина  . В некоторых задачах  знак коэффициента корреляции несуществен и зависит лишь от  выбора единицы измерения. В этом случае в качестве меры различия  между характеристиками используется .

**Методы кластерного анализа.** Методы кластеризации могут быть иерархическими и неиерархическими. Иерархическая кластеризация (hierarchical clustering) характеризуется построением иерархической, или древовидной, структуры.

Иерархические методы могут быть агломеративными (объединительными) и дивизивными. Агломеративная кластеризация (agglomerative clustering) начинается с каждого объекта в отдельном кластере. Кластеры объединяют, группируя объекты каждый раз во все более и более крупные кластеры. Этот процесс продолжают до тех пор, пока все объекты не станут членами одного единственного кластера.

Разделяющая, или дивизивная, кластеризация (divisive clustering) начинается со всех объектов, сгруппированных в единственном кластере. Кластеры делят (расщепляют) до тех пор, пока каждый объект не окажется в отдельном кластере.

Обычно в маркетинговых исследованиях используют агломеративные методы, например методы связи, дисперсионные и центроидные методы. Методы связи (linkage methods) включают метод одиночной связи, метод полной связи и метод средней связи.

В основе **метода одиночной связи** (single method) лежит минимальное расстояние, или правило ближайшего соседа. При формировании кластера первыми объединяют два объекта, расстояние между которыми минимально. Далее определяют следующее по величине самое короткое расстояние, и в кластер с первыми двумя объектами вводят третий объект. На каждой стадии расстояние между двумя кластерами представляет собой расстояние между их ближайшими точками.

**Метод полных связей.** Суть данного метода в том, что два объекта, принадлежащих одной и той же группе (кластеру), имеют коэффициент сходства, который меньше некоторого порогового значения . В терминах евклидова расстояния это означает, что расстояние между двумя точками (объектами) кластера не должно превышать некоторого порогового значения . Таким образом, определяет максимально допустимый диаметр подмножества, образующего кластер.

Метод **средней связи** (average linkage) действует аналогично. Однако в этом методе расстояние между двумя кластерами определяют как среднее значение всех расстояний, измеренных между объектами двух кластеров, при этом в каждую пару входят объекты из разных кластеров.

**Метод Варда.** В этом методе в качестве целевой функции применяют внутригрупповую сумму квадратов отклонений, которая есть ни что иное, как сумма квадратов расстояний между каждой точкой (объектом) и средней по кластеру, содержащему этот объект. На каждом шаге объединяются такие два кластера, которые приводят к минимальному увеличению целевой функции, т.е. внутригрупповой суммы квадратов. Этот метод направлен на объединение близко расположенных кластеров.

В **центроидных** методах (centroid method) расстояние между двумя кластерами представляет собой расстояние между их центроидами (средними для всех переменных). **Центроидный метод (centroid method)** Дисперсионный метод иерархической кластеризации, в котором расстояние между двумя кластерами представляет собой расстояние между их центроидами (средними для всех переменных). Каждый раз объекты группируют и вычисляют новый центроид. Изо всех иерархических методов методы средней связи и Варда показывают наилучшие результаты по сравнению с другими методами.

К другому типу процедур кластеризации относятся **неиерахические методы кластеризации** (nonhierarchical clustering), часто называемые методом -средних. Эти методы включают последовательный пороговый метод, параллельный пороговый метод и оптимизирующее распределение. В последовательном пороговом методе (sequential threshold method) выбирают центр кластера и все объекты, находящиеся в пределах заданного от центра порогового значения, группируют вместе. Затем выбирают новый кластерный центр, и процесс повторяют для несгруппированных точек. После того как объект помещен в кластер с этим новым центром, его уже не рассматривают как объект для дальнейшей кластеризации.

Аналогично работает **параллельный пороговый метод** (parallel threshold method), за исключением того, что одновременно выбирают несколько кластерных центров и объекты в пределах порогового уровня группируют с ближайшим центром.

Метод **оптимизирующего распределения** (optimizing partitioning method) отличается от двух изложенных выше пороговых методов тем, что объекты можно впоследствии поставить в соответствие другим кластерам, чтобы оптимизировать суммарный критерий, такой как среднее внутри кластерное расстояние для данного числа кластеров.

Недостатки неиерархических методов состоят в том, что число кластеров определяется заранее и выбор кластерных центров происходит независимо. Результаты кластеризации могут зависеть от выбранных центров. Многие неиерархические процедуры выбирают первые случаев *( –* число кластеров), не пропуская никаких значений в качестве начальных кластерных центров. Таким образом, результаты кластеризации зависят от порядка наблюдений в данных.

Неиерархическая кластеризация быстрее иерархических методов, и ее выгодно использовать при большом числе объектов или наблюдений.

Возможно использовать иерархические и неиерархические методы в тандеме. Во-первых, первоначальное решение по кластеризации получают, используя такие иерархические методы, как метод средней связи или метод Варда. Полученное этими методами число кластеров и кластерных центроидов используют в качестве исходных данных в методе оптимизирующего распределения.

**Данные.** Кластерный анализ можно применять к интервальным данным, частотам, бинарными данным. Важно, чтобы переменные изменялись в сравнимых шкалах.

Неоднородность единиц измерения и вытекающая отсюда невозможность обоснованного выражения значений различных показателей в одном масштабе приводит к тому, что величина расстояний между точками, отражающими положение объектов в пространстве их свойств, оказывается зависящей от произвольно избираемого масштаба. Чтобы устранить неоднородность измерения исходных данных, все их значения предварительно нормируются, т.е. выражаются через отношение этих значений  к некоторой величине, отражающей определенные свойства данного показателя. Нормирование исходных данных для кластерного анализа иногда проводится посредством деления исходных величин на среднеквадратичное отклонение соответствующих показателей.

Другой способ сводиться к вычислению, так называемого, стандартизованного вклада. Его еще называют -вкладом. -вклад показывает, сколько стандартных отклонений отделяет данное наблюдение от среднего значения:

,

где – значение данного наблюдения, – среднее, – стандартное отклонение. Среднее для -вкладов является нулевым и стандартное отклонение равно 1.

Стандартизация позволяет сравнивать наблюдения из различных распределений. Если распределение переменной является нормальным (или близким к нормальному), и средняя и дисперсия известны или оцениваются по большим выборным, то -вклад для наблюдения обеспечивает более специфическую информацию о его расположении.

Заметим, что методы нормирования означают признание всех признаков равноценными с точки зрения выяснения сходства рассматриваемых объектов. Уже отмечалось, что применительно к экономике признание равноценности различных показателей оправдано не всегда. Желательно наряду с нормированием придать каждому из показателей вес, отражающий его значимость в ходе установления сходств и различий объектов.

Экспертные оценки дают известное основание для определения важности индикаторов, входящих в ту или иную группу показателей. Умножение нормированных значений показателей на коэффициент, соответствующий среднему баллу оценки, позволяет рассчитывать расстояния между точками с учетом неодинакового веса их признаков.

Довольно часто при решении  подобных задач используют не один, а два расчета: первый, в котором все признаки считаются равнозначными, второй, где им придаются различные веса в соответствии со средними значениями экспертных оценок.

**Принятие решения о количестве кластеров.** Главный вопрос кластерного анализа – вопрос о количестве кластеров. Здесь нет твердых правил, позволяющих быстро принять решение, но можно руководствоваться следующим.

При определении количества кластеров руководствуются теоретическими и практическими соображениями. Например, если цель кластеризации – выявление сегментов рынка, то менеджмент может захотеть получить конкретное число кластеров.

В иерархической кластеризации в качестве критерия можно использовать расстояния, при которых объединяют кластеры.

В неиерархической кластеризации чертят график зависимости отношения суммарной внутригрупповой дисперсии к межгрупповой дисперсии от числа кластеров. Точка, в которой наблюдается изгиб или резкий поворот, указывает на приемлемое количество кластеров. Увеличение числа кластеров за эту точку обычно безрезультативно.

Относительные размеры кластеров должны быть достаточно выразительными.

## § 1.3. Введение в теорию нечетких множеств

Пусть - универсальное множество, - элемент , а - определенное свойство. Обычное (четкое) подмножество универсального множества , элементы которого удовлетворяют свойство , определяется как множество упорядоченной пары , где – характеристическая функция, принимающая значение 1, когда x удовлетворяет свойство , и 0 – в другом случае (см. [2], [3], [5], [7], [13], [21], [23] и др.).

Нечеткое подмножество отличается от обычного тем, что для элементов из нет однозначного ответа «нет» относительно свойства . В связи с этим, нечеткое подмножество универсального множества E определяется как множество упорядоченной пари , где – характеристическая функция принадлежности (или просто функция принадлежности), принимающая значение в некотором упорядоченном множестве (например, ).

Функция принадлежности указывает степень (или уровень) принадлежности элемента к подмножеству . Множество называют множеством принадлежностей. Если , тогда нечеткое подмножество может рассматриваться как обычное или четкое множество.

Пример 1. Рассмотрим множество всех чисел от 0 до 10. Определим подмножество множества всех действительных чисел от 5 до 8, .

Покажем функцию принадлежности множества , эта функция ставит в соответствие число 1 или 0 каждому элементу в , в зависимости от того, принадлежит данный элемент подмножеству или нет. Результат представлен на следующем рисунке:

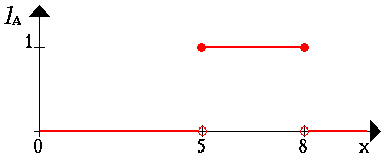


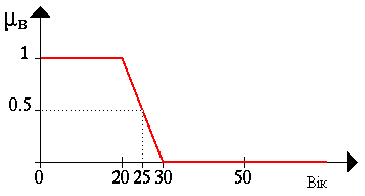
Рис. 1. Функция принадлежности множества A,

Можно интерпретировать элементы, соответствующие 1, как элементы, находящиеся в множестве , а элементы, соответствующие 0, как элементы, не находящиеся в множестве .

Эта концепция используется во многих областях. Но существуют ситуации, в которых данной концепции будет не хватать гибкости.

Пример 2. .

В первом примере мы кодировали все элементы множества с помощью 0 ли 1. Простым способом обобщить данную концепцию является введение значений характеристической функции в единичном интервале .



Мол.люд.

Рис. 2. Характеристическая функция множества молодых людей

На рис. 2 приведена характеристическая функция множества молодых людей.

Пример 3. Пусть , ; – нечеткое множество, для которого ; ; ; ; .

Тогда A можно представить в виде:

или ,

(знак «+» является операцией не сложения, а объединения) или

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  | 0,3 | 0 | 1 | 0,5 | 0,9 |

Пример 4. Пусть , . Нечеткое множество «несколько» можно определить таким образом:

.

Пример 5. Пусть . Нечеткое множество «малый» можно определить:

 .

Пример 6. Пусть

– множество марок автомобилей, а – универсальное множество «стоимость», тогда на мы можем определить нечеткие множества типа: «для небогатых», «для среднего класса», «престижные», с функциями принадлежности типа:



для небогатых для среднего класса престижные

стоимость

Рис. 3. Нечеткие множества: «для небогатых», «для среднего класса», «престижные» с функциями принадлежности

В приведенных выше примерах использованы прямые методы, когда эксперт задает для любого значение  или определяет функцию принадлежности. Как правило, прямые методы задания функции принадлежности используются для измеримых понятий, таких как скорость, час, расстояние, давление, температура и т.д.

Во многих задачах при характеристике объекта можно выделить набор признаков и для любого из них определить полярные значения, отвечающие значениям функции принадлежности, 0 или 1.

Пример 7. В задаче распознавания лица можно выделить следующие признаки:

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | 0 | 1 |
|  | высота лба | низкий | широкий |
|  | профиль носа | курносый | горбатый |
|  | длина носа | короткий | длинный |
|  | разрез глаз | узкий | широкий |
|  | цвет глаз | светлый | Темный |
| … |  |  |  |

Для конкретного лица А эксперт, исходя из приведенной шкалы, задает , формируя векторную функцию принадлежности {A(x1).

Косвенные методы определения значений функции принадлежности используются в случаях, когда нет элементарных измеримых свойств для определения нечеткого множества. Как правило, это методы попарных сравнений. Если бы значение функций принадлежности были известны, например, , , тогда попарные сравнения можно представить матрицей отношений , где (операция деления).

**Операции над нечеткими множествами.** Пусть и – нечеткие множества на универсальном множестве .

– Говорят, что содержится в , если   . Обозначение: . Иногда используют термин «доминирование», то есть в случае если , говорят, что доминирует .

– и равны, если   . Обозначение: .

– Пусть , и – нечеткие множества, заданные на . и дополняют друг друга, если

 – . Обозначение:  или .

Очевидно, что . (Дополнение определено для , но очевидно, что его можно определить для любого упорядоченного ).

– Пересечение: – наибольшее нечеткое подмножество, которое содержится одновременно в и . ).

– Объединение: – наименьшее нечеткое подмножество, которое включает как , так и , с функцией принадлежности:

).

– Разность:  с функцией принадлежности:

).

– Дизъюнктивная сумма: с функцией принадлежности:

.

## § 1.4. Нечеткие системы кластеризации

Взаимосвязь между кластерным анализом и теорией нечетких множеств основана на том обстоятельстве, что при решении задач структуризации сложных систем большинство формируемых классов объектов размыты по своей природе. Эта размытость состоит в том, что переход от принадлежности к непринадлежности элементов к данным классам скорее постепенен, чем скачкообразен. Требуется ответ не на вопрос «Принадлежит ли рассматриваемый элемент тому или иному классу или нет?», а на вопрос: «В какой степени данный элемент принадлежит рассматриваемому классу?»[14].

В общем случае задачей нечеткой кластеризации является нахождение нечеткого разбиения или нечеткого покрытия множества элементов исследуемой совокупности, которые образуют структуру нечетких кластеров, присутствующих в рассматриваемых данных. Эта задача сводится к нахождению степеней принадлежности элементов искомым нечетким кластерам, которые в совокупности и определяют нечеткое разбиение или нечеткое покрытие исходного множества рассматриваемых элементов.

**Общая формальная постановка задачи нечеткого кластерного анализа.** Пусть исследуемая совокупность данных представляет собой конечное множество элементов , которое получило название множество объектов кластеризации. В рассмотрение также вводится конечное множество признаков , каждый из которых количественно представляет некоторое свойство или характеристику элементов рассматриваемой проблемной области. При этом –общее количество объектов, а – количество признаков объектов [14].

Предполагается, что для каждого из объектов кластеризации измерены все признаки множества в некоторой количественной шкале. Каждому из элементов поставлен в соответствие вектор , где – количественное значение признака для объекта данных . Для определенности будем полагать, что все принимают некоторые действительные значения, т.е. .

Процесс измерения свойств может быть реализован в различных шкалах, каждая из которых характеризуется допустимым преобразованием данных. Само множество признаков следует выбирать таким образом, чтобы все признаки были измерены в шкале отношений или шкале интервалов. В этом случае результаты нечеткой кластеризации имеют содержательную интерпретацию, адекватную проблеме нахождения нечетких кластеров [14].

Векторы удобно представить в виде матрицы данных размерности , каждая строка которой равна значению вектора.

Задача нечеткого кластерного анализа формулируется следующим образом: на основе исходных данных D определить такое нечеткое разбиение множества на заданное число c нечетких кластеров , которое доставляет экстремум некоторой целевой функции среди всех нечетких разбиений.

**Уточненная постановка задачи нечеткой кластеризации.** Один из вариантов конкретизации задачи нечеткого кластерного анализа основан на алгоритме ее решения методом нечетких -средних.

Дадим определение нечеткому покрытию нечеткого множества. Система нечетких подмножеств нечеткого множества называется нечетким покрытием, если выполняется следующее условие:

(1)

т.е. объединение всех подмножеств из совпадает (или «покрывает») с исходным нечетким множеством .

Для уточнения вида целевой функции в рассмотрение вводятся некоторые дополнительные понятия. Предполагается, что искомые нечеткие кластеры представляют собой нечеткие множества , образующие нечеткое покрытие исходного множества , при этом условие (1) принимает вид:

(2)

где – общее количество нечетких кластеров , которое считается предварительно заданным и [14].

Далее для каждого нечеткого кластера вводятся в рассмотрение центры искомых нечетких кластеров , которые рассчитываются для каждого из нечетких кластеров по следующей формуле:

 (3)

где m-некоторый параметр, называемый экспоненциальным весом и равный некоторому действительному числу (m > 1). Каждый из центров кластеров представляет собой вектор  в некотором q-мерном нормированном пространстве [14].

В качестве целевой функции будем рассматривать сумму квадратов взвешенных отклонений координат объектов кластеризации от центров искомых нечетких кластеров:

 (4)

где m-экспоненциальный вес нечеткой кластеризации , значение которого задается в зависимости от количества элементов множества . Чем больше элементов содержит множество , тем меньшее значение выбирается для .

Задача нечеткой кластеризации может быть сформулирована следующим образом: для заданных матрицы данных , количества нечетких кластеров и , параметра определить матрицу значений функций принадлежности объектов кластеризации нечетким кластерам , которые доставляют минимум целевой функции (4) и удовлетворяют ограничениям (2), (3), а также дополнительным ограничениям (5) и (6):

 (5)

 (6)

Условие (5) исключает появление пустых нечетких кластеров в искомой нечеткой кластеризации. Последнее условие (6) имеет формальный характер, поскольку непосредственно следует из определения функции принадлежности нечетких множеств. В этом случае минимизация целевой функции (4) минимизирует отклонение всех объектов кластеризации от центров нечетких кластеров пропорционально значениям функций принадлежности этих объектов соответствующим нечетким кластерам [14].

Недостатком постановки задачи нечеткой кластеризации является необходимость априорного задания общего числа нечетких кластеров.

## § 1.5. Обзор программных продуктов для кластеризации многомерных данных.

Нами изучались интернет-материалы на тему «Кластерный анализ, программы кластерного анализа». При обилии ссылок на справочный материал и использование элементов кластерного анализа в статистических исследованиях в различных областях, где требуется обработка статистических данных, наблюдается дефицит программных продуктов, реализующих методы кластерного анализа. В большинстве ссылок упоминаются следующие пакеты программ:​ «ClusterDelta»,  «Сегментация рынка -​​​​​​​ позиционирование товара»,​ «Statistical Package for the Social Sciences»(SPSS Statistics)», MatLab.

В MathCad и Maple стандартных методов кластеризации не найдено, но в некоторых статьях предлагается реализовать их вручную.

​В пакете программ «ClusterDelta – кластерный график как инструмент анализа рынка» (http://tempofox.com/klasternyj-grafik-kak-instrument-analiza-rynka/) кластерный анализ используется для вычисления точечного объема заданных сделок в ограниченном диапазоне времени и финансов (ценовой бар). Кластер представляет собой «ценовой бар», отображаемый на графике, то есть бар, который разбит на ценовые уровни, на каждом из которых отображается проторгованный объем [26].

Пакет программ  «Сегментация рынка – ​​​​​​​ позиционирование товара» (<http://www.segmentation-market.ru/index.html>) позволяет производить:

1) выделение сегментов рынка методами экспертного сегментирования;

2) выделение целевых сегментов рынка методами кластерного анализа;

3) при оценке сходства потребителей учитывается важность отдельных переменных;

4) анализ перекрытия выделяемых сегментов рынка, сегментация рынка;

5) сегментация рынка - выделение и интерпретация семантики выделенных сегментов;

6) исследование выделенных сегментов с помощью функций полезности, главных факторов, профилей сегментов при сегментации рынка;

7) позиционирование товаров и бренда на выделенных сегментах, построение карты восприятия товара (карта позиционирования);

8) исследование стратегического потенциала целевых сегментов рынка;

9) представление результатов анализа доходных сегментов рынка в виде таблиц и графиков.

Используемый в этом программном пакете метод K-means – это итерационный метод кластеризации. Метод разбивает множество объектов на заданное аналитиком число кластеров (K). Объединение объектов в кластеры происходит на основе вычисляемой меры сходства. После первого назначения объектов кластерам (первой итерации) происходит повторное вычисление мер сходства и перенос объектов из одного кластера в другой. Итерационный процесс продолжается до тех пор, пока не будет получено наилучшее разбиение объектов на кластеры или достигнуто максимальное число итераций [27].

Программный пакет​ «StatisticalPackagefortheSocialSciences» (SPSSStatistics) http://www.datuapstrade.lv/rus/spss/section\_20/2/ – «статистический пакет для социальных наук» –  компьютерная программа для [статистической обработки](https://docviewer.yandex.ru/r.xml?sk=y7641a37c0c52a9f4a3be81098b2d23ee&url=http%3A%2F%2Fru.wikipedia.org%2Fwiki%2F%25D0%25A1%25D1%2582%25D0%25B0%25D1%2582%25D0%25B8%25D1%2581%25D1%2582%25D0%25B8%25D1%2587%25D0%25B5%25D1%2581%25D0%25BA%25D0%25B8%25D0%25B9_%25D0%25B0%25D0%25BD%25D0%25B0%25D0%25BB%25D0%25B8%25D0%25B7%22+%5Co+%22%D0%A1%D1%82%D0%B0%D1%82%D0%B8%D1%81%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D0%B9+%D0%B0%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7) данных, один из лидеров рынка в области коммерческих статистических продуктов, предназначенных для проведения прикладных исследований в [социальных науках](https://docviewer.yandex.ru/r.xml?sk=y7641a37c0c52a9f4a3be81098b2d23ee&url=http%3A%2F%2Fru.wikipedia.org%2Fwiki%2F%25D0%25A1%25D0%25BE%25D1%2586%25D0%25B8%25D0%25B0%25D0%25BB%25D1%258C%25D0%25BD%25D1%258B%25D0%25B5_%25D0%25BD%25D0%25B0%25D1%2583%25D0%25BA%25D0%25B8%22+%5Co+%22%D0%A1%D0%BE%D1%86%D0%B8%D0%B0%D0%BB%D1%8C%D0%BD%D1%8B%D0%B5+%D0%BD%D0%B0%D1%83%D0%BA%D0%B8) [7], [28].

Он использует 7 методов кластеризации: межгрупповые связи (Between-groups linkage); связь внутри групп (Within-groups linkage); близлежащий сосед (Nearest neighbor); дальний сосед (Furthest neighbor); центроидная кластеризация (Centroid clustering); медианная кластеризация (Median clustering); метод Варда (Ward's Method). Этот пакет не ориентирован на решение задач кластеризации в условиях неоднородных показателей, на обработку нечетких данных.

В Маtlab [14] реализуются методы четкой кластеризации: метод ближнего соседа, метод дальнего соседа, метод средней связи, центроидный метод, метод медианной связи. Возможна реализация нечеткого fcm-метода с использованием командной строки или графического интерфейса пользователя. Для нахождения центров кластеров в Маtlab имеется встроенная функция fcm: [*center*, *U*, *obj*\_*fcm*] = *fcm*(*data*, *cluster*\_*n*).

Аргументами данной функции являются: data – множество данных, подлежащих кластеризации, каждая строка описывает точку в многомерном пространстве характеристик; cluster\_n – количество кластеров (более одного).

Функцией возвращаются следующие параметры: center – матрица центров кластеров, каждая строка ко­торой содержит координаты центра отдельного кластера; U – результирующая матрица ФП; obj\_fcm – значение целевой функции на каждой итерации.

В «MatLab» [14] имеются встроенные функции, позволяющие реализовать алгоритмы четкой и нечеткой кластеризации. Этот пакет рассчитан на использование профессиональными математиками и программистами и в этом их ограничение.

Таблица со сравнительным анализом программных пакетов для кластеризации приведена в приложении 1.

## Выводы

Кластеризация – это разбиение множества объектов на некоторые однородные подмножества (кластеры), параметры которых изначально известны. Для кластера характерны внутренняя однородность (объекты одного класса схожи между собой по определенным признакам) и внешняя изолированность (объекты разных классов существенно отличаются).

Методы кластеризации делятся на иерархические и неиерархические. В иерархических – происходит последовательное объединение маленьких кластеров в большие или наоборот разделении больших кластеров на маленькие. Неиерархические методы основаны на разделении набора данных на определенное количество кластеров и выполнении итеративного процесса оптимизации некоторой целевой функции, определяющей оптимальность (обусловленную особенностями алгоритма) данного разбиения множества объектов на кластеры, на итеративный процесс накладывается условие остановки.

При решении задач структуризации сложных систем большинство формируемых классов объектов размыты по своей природе. Это приводит к необходимости использования «нечеткой кластеризации». Задачей нечеткой кластеризации является нахождение нечеткого разбиения или нечеткого покрытия множества элементов исследуемой совокупности, которые образуют структуру нечетких кластеров, присутствующих в рассматриваемых данных.

Кластеризация применяется в статистике, экономике, финансовой математике, в социологии, в информатике для «интеллектуального» анализа данных и др.

Изучение возможностей программных пакетов показало, что, с одной стороны, существуют узконаправленные пакеты программ для конкретных предметных областей, например, экономики: для сегментации рынка товаров, построения ценового бара и т.д. С другой стороны профессиональные программные пакеты с обширным набором кластерных методов, но без привязки к терминологии предметных областей. Эти программные пакеты не ориентированы на решение задач кластеризации в условиях неоднородных показателей, на обработку нечетких данных. Рассчитаны на использование профессиональными математиками и в этом их ограничение.

# Глава 2. Методика кластеризации ИТ-предприятий

## 2.1 Выбор критериев и параметров, используемых при кластеризации ИТ-предприятий

Понятие экономического кластера введено М. Портером [17] и обозначает географи­чески сконцентрированные компании, поставщиков услуг, фир­мы в схожих отраслях промышленности, дочерние учреждения (как, например, институты, клубы качества, торговые ассоциа­ции и пр.) в определенной области, которые и конкурируют, и сотрудничают между собой, неизбежно создавая единый про­цесс взаимодействия.

К понятию экономического кластера М. Портер пришел из анализа конкурентоспособности отдель­ных отраслей экономики той или иной страны на мировом рынке. Он показал, что сконцентрированные географически и не жестко конкурирующие между собой взаимосвязанные груп­пы предприятий могут представлять достаточно мощную кон­курентоспособную силу в глобальной экономике.

В литературе достаточно часто в качестве примера обращаются к таким экономическим кластерам, как калифорнийский винный кластер, шведский кластер лесной продукции, кембриджский кластер породистых лошадей и др. Подробно исследованы и описа­ны их состав, географические особенности, причины, по кото­рым данные компании объединены в кластеры, конкурентные преимущества.

Актуальным является исследование текущего уровня ИТ-сектора Ульяновской области, применение методов кластерного анализа, сравнение теоретического анализа данных с намечающимися кластерами ИТ-предприятий.

Исследование ИТ-сектора – процесс сложный и неоднозначный, единой методики и показателей в аналитических материалах анализа ИТ-сектора не существует, однако есть пересечения, в основном по показателям официальной статистики. При проведении исследований используются различные подходы и даже разные определения ИТ-отрасли. Возникают проблемы выбора критериев кластеризации ИТ-предприятий и соответствующих им параметров.

Приведем краткий обзор статистических исследований и данных такого рода [22]. Аналитические материалы профессиональных и академических (исследовательских) сообществ с статистическими исследованиями ИТ-сектора представлены:

1) в отчетах государственной статистики и на основе исследований государственных учреждений ([www.gks.ru/bgd/regl/b12\_14p/Main.htm](http://www.gks.ru/bgd/regl/b12_14p/Main.htm), <http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/science_and_innovations/it_technology/>, http://www.hse.ru/primarydata/rir2012ФГУП ВНИИ ПВТИ) <http://pvti.ru/>) [29]-[31];

2) в аналитических материалах и исследованиях профессиональных сообществ и интернет-изданий [АП КИТ http://www.apkit.ru/files/Strategy\_APKIT\_2012\_vr.pdf, РУССОФТ, http://www.russoft.ru/reports, Институт развития информационного общества (ИРИО) http://eregion.ru/, TAdviser <http://www.tadviser.ru/>) [32]-[41];

3) в отчетах рейтинговых агентств (ЭКСПЕРТ РА <http://www.raexpert.ru/>, CNEWS <http://www.cnews.ru/reviews/new/rynok_it_itogi_2012/>, http://www.ratingruneta.ru/, <http://www.cmsmagazine.ru/>, http://raec.ru/analytics/, <http://2012.tagline.ru/>) [42]-[49].

Кратко охарактеризуем эти источники. Сбалансированная система показателей (Balanced Scorecard), позволяет описать механизм взаимосвязи стратегических замыслов и решений сотрудников с ежедневными задачами и предлагает выбор соответствующих количественных и качественных показателей.

Ежегодные исследования профессиональной ассоциации РУССОФТ представлены в виде обзоров <http://www.russoft.ru/reports> [34]. В обзор 2012 г. включены 120 заполненных анкет и аналитические материалы зарубежных аналитических агентств. Опрос был проведен с использованием технологий компании ComfirmIT с привлечением маркетингового агентства «Той-Опинион». Направления исследований включают показатели в основном по объему и структуре экспорта программного обеспечения (ПО) и позиции РФ на мировом рынке ПО.

Аналитический доклад 2012 г. ассоциации АП КИТ http://www.apkit.ru/files/Strategy\_APKIT\_2012\_vr.pdf, «О мерах по развитию отрасли ИТ в Российской Федерации» включает следующие разделы: перечень секторов ИТ-рынка, мировые тренды развития ИТ и роль отрасли ИТ в развитии российской экономики и приоритетные технологии для российского ИТ-рынка, структура российского ИТ-рынка в сравнении с мировым по объемам реализации и потребителям, а также по кадровому составу (выпускники)[44].

Ресурс Федерального государственного унитарного предприятия «Всероссийский научно-исследовательский институт проблем вычислительной техники и информатизации» (ФГУП ВНИИ ПВТИ) <http://pvti.ru/> публикует материалы по Мониторингу региональной информатизации и прогнозы (73 субъекта РФ) http://pvti.ru/monitoring-materials.htm по заданию федеральных органов государственной власти на основе типового «Паспорта информатизации субъекта Российской Федерации». Ими используется программный продукт «БАРС.Web.Своды» (г.Казань), основанный на веб-технологиях. Система обеспечивает автоматизацию процесса сбора, консолидации и анализа информации, предоставляемой субъектами Российской федерации и характеризующей состояние работ по использованию и развитию информационных технологий на региональном и муниципальном уровне [45].

Сборник публикаций «Индикаторы информационного общества: 2013» http://www.hse.ru/primarydata/iio2013, подготовленный в НИУ ВШЭ, посвящен статистике информационного общества. В нем представлена статистическая информация о деятельности организаций сектора информационных и коммуникационных технологий (ИКТ), об использовании ИКТ в организациях (предпринимательского сектора, государственного управления, социальной сферы) и населением, доступе к вычислительной технике, средствам связи и Интернету домашних хозяйств. Впервые приведены данные конъюнктурного обследования деловой активности сектора услуг, включающие оценки состояния и тенденций динамики важнейших показателей деятельности организаций, оказывающих ИКТ-услуги [46].

Интерес представляет данные и рейтинги, например, представленные на ресурсе CNEWS <http://www.cnews.ru/reviews/new/rynok_it_itogi_2012/>, который регулярно формирует рейтинг компаний ИТ-сектора «CNEWS 100», публикует аналитику, при этом основные показатели – это выручка региональных ИТ-компаний, их доля в бюджетах, распределение количества компаний по регионам и динамика за последний год [38]. В основном в рейтинг вошли Московский регион, Ульяновских компаний – нет.

Анализ присутствия региональных ИТ-компаний в рейтингах, приведенных на ресурсах <http://2012.tagline.ru/top135digital/>, http://www.ratingruneta.ru/research/web2010, http://www.raexpert.ru/ невелико [37], [42], [47]. В основном первые строчки занимают компании Москвы и Санкт-Петербурга. Количество компаний в рейтингах (обычно из 400 компаний).

На Интернет-ресурсе в разделе «Показатели развития информационного общества» приведены 122 интегральных показателя <http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/science_and_innovations/it_technology/> [49]. Эти показатели классифицированы на две большие группы: факторы развития и использование ИКТ для развития. Внутри каждой группы приведена более детальная классификация.

Анализ приведенных группировок показателей и характеристик ИКТ-уровня субъектов РФ показывает, что, во-первых, в рамках официальных статистических данных в основном присутствует взгляд на ИКТ с позиции отрасли связи и телекоммуникации, во-вторых, в новых данных присутствует уклон в сторону оценивания ИКТ, как инновационной деятельности.

В монографии Н.Г. Ярушкиной, Т.В. Афанасьевой и др. «Исследование ИТ-кластера Ульяновской области» [22] приведена классификация характеристик, отражающая основные компоненты и уровень ИТ-кластера, предложена следующая классификация основных характеристик ИТ-кластеров:

a. инфраструктура и использование ИТ, так как именно эти показатели являются фундаментом в развитии ИТ- кластера;

b. финансовые показатели ИТ-кластера и влияние развития ИТ на экономическое состояние, так как в этом направлении существует значительная региональная дифференциация, при этом и экономическое состояние влияет на развитие ИТ-кластера;

c. показатели, характеризующие участников рынка ИТ-сектора – пользователи ИТ-продуктов и производители ИТ-продуктов,- именно они определяют спрос и предложение;

d. используемые технологии, так как это позволит выделить конкурентоспособные технологии, присущие определенному кластеру, а также сопоставить их с основными трендами;

e. инновации и исследования в ИТ, так как они определяют направления развития ИТ [22].

Эта классификация согласуется с классификациями, приведенными в рассмотренных выше интернет-ресурсах, является более грубой, с одной стороны, а с другой стороны, позволяет в дальнейшем группировать показатели из различных источников.

В монографии Н.Г. Ярушкиной, Т.В. Афанасьевой и др. «Исследование ИТ-кластера Ульяновской области» приведена анкета для анализа деятельности и рынков ИТ-предприятий [22].

Для обработки результатов анкетирования, в силу многочисленности параметров и их различной природы, вопросы анкет разделены нами на следующие группы, что базируется на приведенной выше группировке. Ниже приведены группы и отвечающие им параметры.

1. Область деятельности компании.

1.1. Область деятельности компании (термины).

2. Финансово-экономические показатели.

2.1. Объем реализации IT-продуктов/услуг за год (рубли).

2.2.Объем реализации IT-продуктов/услуг за год в других странах (руб.)

2.3. Прирост объема реализации (прибыли) за год (%).

2.4. Доля продаж в другие страны (%).

2.5. Доля  внешнего финансирования (%).

2.6. Доля самофинансирования (%).

2.7. Доля заемных средств (кредиты) (%).

2.8. Доля финансирования по госзаказам (%).  
3. Рынок сбыта.

3.1. Количество клиентов (число).

3.2. Количество выполненных проектов (число).

3.3. Область распространения продукции/услуг (страны, города, термины).

3.4. Конкуренты, присутствующие на рынке, и их продукция (термины).

3.5. Рейтинги, позволяющие выявить конкурентов (термины).

3.6. Конкурентные преимущества, недостатки и особенности вашей продукции (термины).

3.7. Уровень удовлетворенности потребителя(%).

3.8. Наиболее значимые награды вашей компании (термины).

3.9. Оцениваемая емкость рынка (термины, числа).

3.10. Доля рынка, занимаемая компанией (%).

3.11. Средняя цена вашей продукции (услуг) (рублей / 1 час, рубль).

4. Кадровые показатели.

4.1. Средний возраст сотрудников (возраст, числа).

4.2. Сотрудников с высшим образованием (%).

4.3. Сотрудников, являющихся студентами (%).

4.4. Количество и сотрудников с научными степенями (%).

4.5. Сотрудников, являющихся выпускниками 2013 года (%).

4.6. Есть ли корпоративное обучение иностранным языкам (да, нет).

4.7. Есть ли корпоративное обучение новым технологиям разработки ПО, маркетингу и т.п. (да, нет).

4.8. Количество сотрудников (за последние 5лет), участвующих в конференциях,  семинарах, выступающих с лекциями (числа).

4.9. Количество сотрудников (числа).

5. Внутренние процессы компании.

5.1. Стратегия развития компании (термины).

5.2. Возраст компании (числа).

5.3. Перечислите технологии и инструменты, используемые для разработки продуктов (реализации услуг) (термины).

5.4. Перечислите технологии и инструменты, используемые для реализации внутренних процессов компании (термины).

5.5. Есть ли сертификация по ISO серии 9000 (да, нет).

5.6. Оцениваемый уровень зрелости процессов компании (термины).

5.7. Наличие отдела маркетинга (да, нет).

5.8. Наличие отдела продаж (да, нет).

6. Инновационная деятельность.

6.1. Доля инновационных наукоемких продуктов в объеме реализации (%).

6.2. Количество договоров на выполнение НИОКР с НИИ или ВУЗами (числа).  
7. Социальные программы и образовательная деятельность.

7.1.Объем финансирования социальных программ в % от объема реализации (%).

7.2. В организации каких олимпиад и конкурсов принимала участие ваша компания (термины).

7.3. В организации каких лекций, семинаров и конференций принимала участие ваша компания (термины)?

7.4. В организации каких летних, воскресных  школ, лагерей принимала участие ваша компания (термины)?

7.5. Среднее количество студентов, проходящих производственную практику на  базе Вашего предприятия (числа).

7.6. Количество стипендий от Вашей организации для талантливой молодежи (числа).

7.7. Количество грантов для талантливой молодежи от Вашей организации (числа, термины).

7.8. Количество разработанных профессиональных тестов(числа, термины).

7.9. Количество учебных программ (специалистов, бакалавров, магистров),  созданных при участии Ваших сотрудников (числа, термины).

7.10. Количество рабочих программ дисциплин, созданных при участии Ваших  сотрудников(числа, термины).

7.11. Количество дисциплин, в преподавании которых участвуют сотрудники  в образовательных заведениях (вузы, ссузы, школы, дошкольные   учреждения) (числа, термины).

## § 2.2. Анализ и предварительная обработка данных анкет руководителей ИТ-предприятий г. Ульяновска

Коллективом ученых УлГТУ (Н.Г. Ярушкина, Т.В. Афанасьева, О.В. Шиняева, К.В. Святов, Л.М. Валкин, Д.А. Ефремов, К.Г. Калимуллин и др.) в рамках проекта «Исследование ИТ-кластера Ульяновской области» в 2013г. проведено анкетирование руководителей ИТ-предприятий г. Ульяновска. Собраны данные по 87 предприятиям по 39 вопросам анкеты (см. [22 с.112 ]).

При анализе и обработке анкетных данных возникли следующие проблемы. Во-первых, кластеризация требует объединения близких по смысловому значению параметров в группы и сравнения объектов (ИТ-предприятий) по этим группам параметров (см. § 2.1).

Во-вторых, анкетные данные разнородны. Среди них числа (рубли, количество сотрудников и т.д.), проценты, термины. Требуется перевод лингвистических переменных в числовые, требуется приведение всех данных к одному типу единиц измерения.

В-третьих, велик разброс данных в пределах одной группы. Теория кластерного анализа показывает, что в этом случае кластеризация затруднительна. Требуется нормировка данных.

В-четвертых, часть вопросов анкеты руководители предприятий оставили без ответа. Предприятия имеют разный уровень развития и данные некоторых предприятий составляют долю, близкую к 0, по сравнению с лидерами. Это создает трудности при обработке данных. Матрица 87 предприятий на 39 ответов имеет значительные пустующие поля. Требуется алгоритм уменьшения размерности матрицы, исключения из рассмотрения предприятий и вопросов анкеты с нулевыми или малыми значениями параметров.

Предварительная обработка данных в нашем случае включает в себя:

– приведение данных к однородному виду:

– ( Uniformity);

– анализ;

– редукцию (редукция в технике – сокращение, уменьшение силы движения, напряжения): – ( Reduction).

Предварительная обработка данных производилась в указанном порядке: .

Входные данные программы можно представить в виде таблицы 1. По горизонтали приведена нумерация параметров согласно списку из § 2.1.

**Таблица 1.**

**Результаты анкетирования руководителей ИТ-предприятий Ульяновской области за 2012г. (87 предприятий, 39 вопросов)**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Группы** | **Область деятельности компании** | **Финансово-экономические показатели** | | | | | | | |
|  | **Название компании** | **1.1** | **2.1 (руб)** | **2.2(руб)** | **2.3** | **2.4** | **2.5** | **2.6** | **2.7** | **2.8** |
| 1 | ИнтелСофт | Разработка собственных программных продуктов, Обучение IT-технологиям | 0 - 500 тыс | 0 |  |  | 0 | 100 | 0 | 0 |
| 2 | СимбирСофт |  |  |  | 50 |  | 100 |  |  |
| 3 | ИП Хомяченко | 0 - 500 тыс. |  |  |  |  | 100 |  |  |
| 4 | JackNyfe (Echo) |  |  |  |  |  |  |  |  |
| 5 | Группа Компаний ИТМ | 10 - 50 млн. | 0 - 500 тыс. |  |  |  |  |  |  |
| 6 | ООО "Купи батон" | 1 - 10 млн. | 1 - 10 млн. | 0,7 | 0,8 | 0 | 100 | 0 | 0 |
| 7 | JackNyfe(Echo) | 100 млн. | 50 - 100 млн. |  | 100 | 0 | 100 | 0 | 0 |
| 8 | ООО АИСТ | 1000 тыс |  |  |  | 0 | 100 | 0 | 0 |
| 9 | ООО"Телеком.ру" |  |  |  |  |  |  |  |  |
| 10 | ITmaster | 1000 тыс |  | 20 |  |  | 100 |  |  |
| 11 | Effective Script |  |  |  |  |  |  |  |  |
| 12 | Веб-студия Creater | 1 - 10 млн. | 1 - 10 млн. | 50 |  | 0 | 100 | 0 | 0 |
| 13 | ооо "андер девелопмент" |  |  |  |  |  |  |  |  |
| 14 | ITECH.group | - 500 млн. |  | 40 |  |  |  |  |  |
| 15 | ООО "АктивУчет" | 0 - 500 тыс. |  | - 7 % |  |  |  |  |  |
| 16 | ООО «Решение» | 1 - 10 млн. | 0 - 500 тыс. | 5 | 1 |  | 100 |  |  |
| 17 | ООО «Тауруна» | 1 - 10 млн. |  |  |  |  | 100 |  |  |
| 18 | ООО "УмКо" | 1 - 10 млн. |  |  |  |  | 100 |  |  |

Продолжение таблицы 1.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Рынок сбыта** | | | | | **Кадры** | | |
|  | **3.1** | **3.2** | **3.3** | **3.7** | **3.11** | **4.2** | **4.4** | **4.5** |
| 1 | 1 | 2 | г. Ульяновск | 0,75 | Средняя стоимость человеко-часа у нас - 650р./ч. | 75 | 25 | 4 |
| 2 |  |  |  |  |  | 95 | 7 7% | 70 |
| 3 |  |  |  |  |  | 100 |  | 1 |
| 4 |  |  |  |  |  | 90 |  |  |
| 5 |  |  |  |  |  | 80 | 0 | 28 |
| 6 | 2898365 | 1 | Весь мир | 0,9 | Затраты на разработку программного комплекса составят 4 миллиона рублей. | 90 | 5 | 10 |
| 7 |  |  |  |  |  | 97 | 1 чел 3 % | 30 |
| 8 |  | 3 |  |  |  |  |  | 3 |
| 9 |  |  |  |  |  | 90 | 0 |  |
| 10 | 90 |  |  | 1/50 | 600 рублей в час | 90 | 0 | 11 |
| 11 |  |  |  |  |  |  |  |  |
| 12 | 100 | 150 | Россия, США | 0.94 | 1000 рублей / 1 час | 90 | 0 | 16 |
| 13 |  |  |  |  |  | 80 |  | 30 |
| 14 | 214 | 142 |  | 100 | 2000 руб/человекочас | 83 | 2,3 | 119 |
| 15 | 21 |  | Димитровград |  | 800 руб. час | 100 % |  | 1 |
| 16 | 124 | 35 | Москва, Петербург, Карелия, Новосибирск, Алтай, Магадан, Воронеж. |  | 700 рублей | 100 |  | 1 |
| 17 |  |  |  |  |  | 100 |  | 8 |
| 18 |  |  |  |  |  | 100 |  | 8 |

Продолжение таблицы 1.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Внутренние процессы компании** | | **Инновационная деятельность** | **Соц программы** | |
| 1 | **5.7** | **5.8** | **6.1** | **7.6** | **7.7** |
| 2 |  |  | 30 | 0 | 0 |
| 3 | 1 | 5 |  | 0 | 0 |
| 4 |  |  | 100 |  |  |
| 5 | в США | в США |  |  |  |
| 6 | 7 | 2 | 0 | 0 | 0 |
| 7 | 5 | 2 | 100 | - | - |
| 8 | есть в США, 1 чел | есть в США , штатных сотрудников 2 + фрилансе количество меняется |  | 0 | 0 |
| 9 | 0 | 0 |  |  |  |
| 10 | 5 | 10 |  |  |  |
| 11 | 1 | 2 | 0 |  |  |
| 12 |  |  |  |  |  |
| 13 | 1 | 2 | 0 | 0 | 0 |
| 14 |  |  |  |  |  |
| 15 | 3 сотрудника | 5 сотрудников |  |  |  |
| 16 | нет |  |  |  |  |
| 17 | В отделе маркетинга работают 1 сотрудника. | В отделе продаж работают 1 сотрудника. | 10 |  |  |
| 18 |  |  |  |  |  |

Нами обрабатывались анкетные данные 87 ИТ-предприятий по 39 параметрам. Т.е. исходная матрица данных имела размерность 87×39. В таблице 1 приведены данные по нескольким предприятиям и по части параметров. Полный вариант таблицы 1 см. на лазерном диске.

**Входные данные программы.** Этой таблице поставим в соответствие матрицу, элементами которой являются данные из таблицы. Далее будем говорить о матрице и её элементах .

Данные в таблице 1 представлены как числовыми значениями, так и другими типами данных, как слова, текст и др., поэтому необходимо приведение данных к одному типу (однородность данных).

**Приведение данных к однородному виду.** Приведение данных к однородному виду производится отдельно в каждой группе.

1) Область деятельности компании.  
2) Финансово-экономические показатели.  
3) Рынок сбыта.  
4) Кадровые показатели.  
5) Внутренние процессы компании.  
6) Инновационная деятельность.  
7) Социальные программы и образовательная деятельность.

Специфика группы параметров «Область деятельности компании» заключается в том, что в группе один параметр, но он текстового типа данных.

**Таблица 2.**

Матрица преобразований и значений

параметра «Область деятельности компании».

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Назва-ние компа-нии | Обла-сть деяте-льно-сти компа-нии | Pазра-ботка | Продажа | Ремонт и обслу-жива-ние, сопро-вождение | Интер-нет и телекоммуникации | Обучение | Науч-ные иссле-дования | Защи-та информации | Значе-ние | Нормирован-ные значе-ния |
| ИнтелСофт | Заказ-ная разра-ботка про-грам-много обеспечения | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0,27128 | 0,444 |
| … | … | … | … | … | … | … | … | … | … | … |
| ФГУП "ЦентрИнформ" | Заказ-ная разра-ботка про-грам-много обеспечения | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0,02128 | 0,035 |
|  | Кол-во | 47 | 36 | 36 | 12 | 4 | 3 | 2 |  |  |

**Таблица 3.**

Общий вид матрицы преобразований и векторов , .

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Назва-ние компа-нии | Обла-сть деяте-льно-сти компа-нии | Pазра-ботка | Продажа | Ремонт и обслу-жива-ние, сопро-вожде-ние | Интер-нет и телекоммуникации | Обучение | Науч-ные иссле-дования | Защита информации | Зна-чение | Нормирован-ные значе-ния |
| ИнтелСофт | text0 | a0,0 | a0,1 | a0,2 | a0,3 | a0,4 | a0,5 | a0,6 | zn0 | nzn0 |
| … | … | … | … | … | … | … | … | … | … | … |
| ФГУП "ЦентрИнформ" | textn | an,0 | an,1 | an,2 | an,3 | an,4 | an,5 | an,6 | znn | nznn |
|  | Кол-Во | 47 | 36 | 36 | 12 | 4 | 3 | 2 |  |  |

Рассмотрим способы заполнения и назначение этих таблиц.  
Дан вектор состоящий из элементов (лингвистические переменные), , – номер предприятия по таблице 1. Необходимо его преобразовать в вектор , состоящий из элементов (числовые дробные значения), .

Преобразование осуществляется с помощью матрицы преобразования , состоящей из элементов (одно из двух числовых значений, либо либо ), j , – номер предприятия по таблице 1, – номера признаков (сфер деятельности), перечисленных по горизонтали в таблице 3.

Матрица преобразования строится по следующему принципу. Выделены несколько основных областей деятельности, которые теперь являются лингвистическими переменными: разработка; продажа; ремонт и обслуживание, сопровождение; интернет и телекоммуникации; обучение; научные исследования; защита информации. Проведён анализ текстовых записей на предмет принадлежности к области деятельности ( – если компания обладает признаком , – если нет).

Вектор значений строится с помощью матрицы преобразования по следующей формуле:

в знаменателе – количество ИТ-предприятий, обладающих признаком ; – весовой коэффициент для признака . Наибольший весовой коэффициент приобретает тот род деятельности, которым занимается меньше всего предприятий. – интегральный параметр, отражающий сферы деятельности -го предприятия, получаемый как сумма параметров строки таблицы 3 с соответствующими весовыми коэффициентами.

В этом случае каждый вектор-столбец таблицы делится на его - норму. Для полученных векторов-строк тоже ищется – норма. Это один из возможных способов ввода весовых коэффициентов. На практике часто для экономических моделей используют экспертные оценки весовых коэффициентов.

После получения вектора значений получим вектор нормированных значений по формуле:

.

Элементы получившегося вектора выражают, какую долю составляет параметр по отношению к максимальному параметру столбца, т.е. долю от показателя предприятия-победителя. Нормировка производится со следующими целями: чтобы все данные при кластеризации имели одинаковые единицы измерения (в долях от параметров-победителей), для уменьшения разброса данных, для приведения значений всех параметров в промежуток от 0 до 1. Здесь для нормирования векторов использована супремум-норма.

Во всех оставшихся группах (финансово-экономические показатели, рынок сбыта, кадровые показатели, внутренние процессы компании, инновационная деятельность, социальные программы и образовательная деятельность) параметры преимущественно числовые: тысячи рублей, проценты, да/нет(/). Их достаточно просто привести к нормированным числовым значениям.

Построим отдельно для каждой группы матрицу значений , состоящую из элементов (числовые дробные значения), ; , – номер предприятия, – номер признака внутри каждой из заявленных выше групп признаков. Перейдем к матрице нормированных значений , состоящей из элементов (числовые дробные значения), ; , вычисляемых по формуле:

.

Элементы получившейся матрицы выражают, какую долю составляет параметр по отношению к максимальному параметру -ого столбца, т.е. долю от аналогичного показателя предприятия-победителя. Нормировка производится со следующими целями: чтобы все данные при кластеризации имели одинаковую размерность (в долях от параметров-победителей), для уменьшения разброса данных, для приведения значений всех параметров в промежуток от 0 до 1. Здесь для нормирования векторов использована супремум-норма.

**Анализ данных.** После приведения данных к однородному виду мы получили матрицу данных . Ее элементами являются параметры дробного вида от до , включительно.

Если значение параметра равно , то будем считать, что параметр в анкете не был заполнен при анкетировании.

Построим матрицу наполненности анкетных данных по формуле:

.

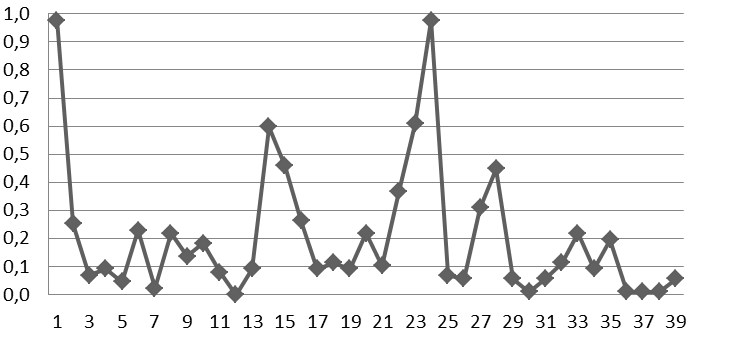
По матрице построим вектор «плотности заполнения анкет по параметрам» , где – номер вопроса анкеты по таблице 1,

где в числителе – количество заполненных предприятиями -ых вопросов анкеты,– плотность заполнения этого пункта анкеты 87 предприятиями, участвующими в анкетировании, в нашем случае

Построим вектор «плотность заполнения анкет по предприятиям» , где – номер предприятия по таблице 1.

в числителе – количество заполненных вопросов анкеты -ым предприятием, – плотность заполнения, в нашем случае .

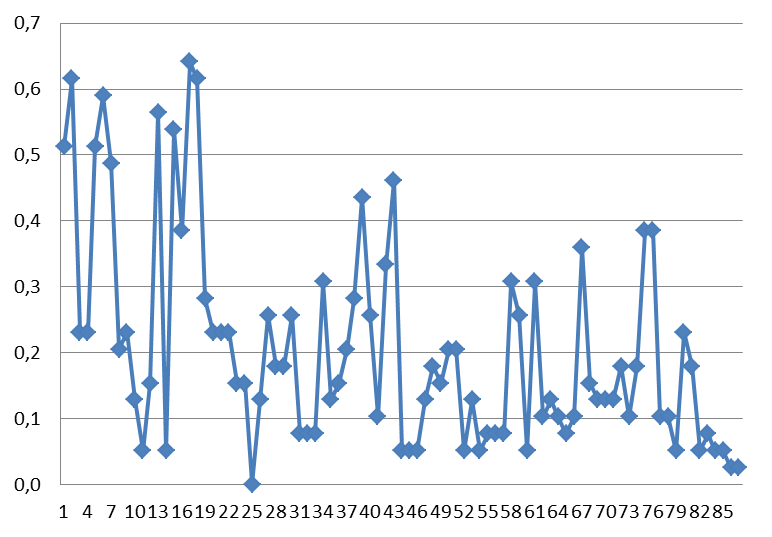
Построим график плотности заполнения 39 пунктов анкеты.



Пункты анкеты

Плотность заполнения пункта анкеты

Рис. 4. График плотности заполнения 39 пунктов анкеты.



Номер предприятия

Плотность заполнения анкеты предприятиями

Рис. . Плотность заполнения пунктов анкеты 87 предприятиями.

Рис. 5. Плотность заполнения анкеты каждым из 87 ИТ-предприятий

Воспользовавшись этими рисунками, можно получить представление о наполняемости анкеты ответами и выбрать уровень плотности заполнения строк и столбцов таблицы данных, выше которого данные можно считать заполненными. Данные ниже выбранного уровня можно обнулить, соответствующие строки – для предприятий, столбцы – для параметров исключим из таблицы значений.

**Редукция данных.** В нашем же случае «редукция» – сокращение объема данных. Редукцию можно провести на основе результатов анализа данных.

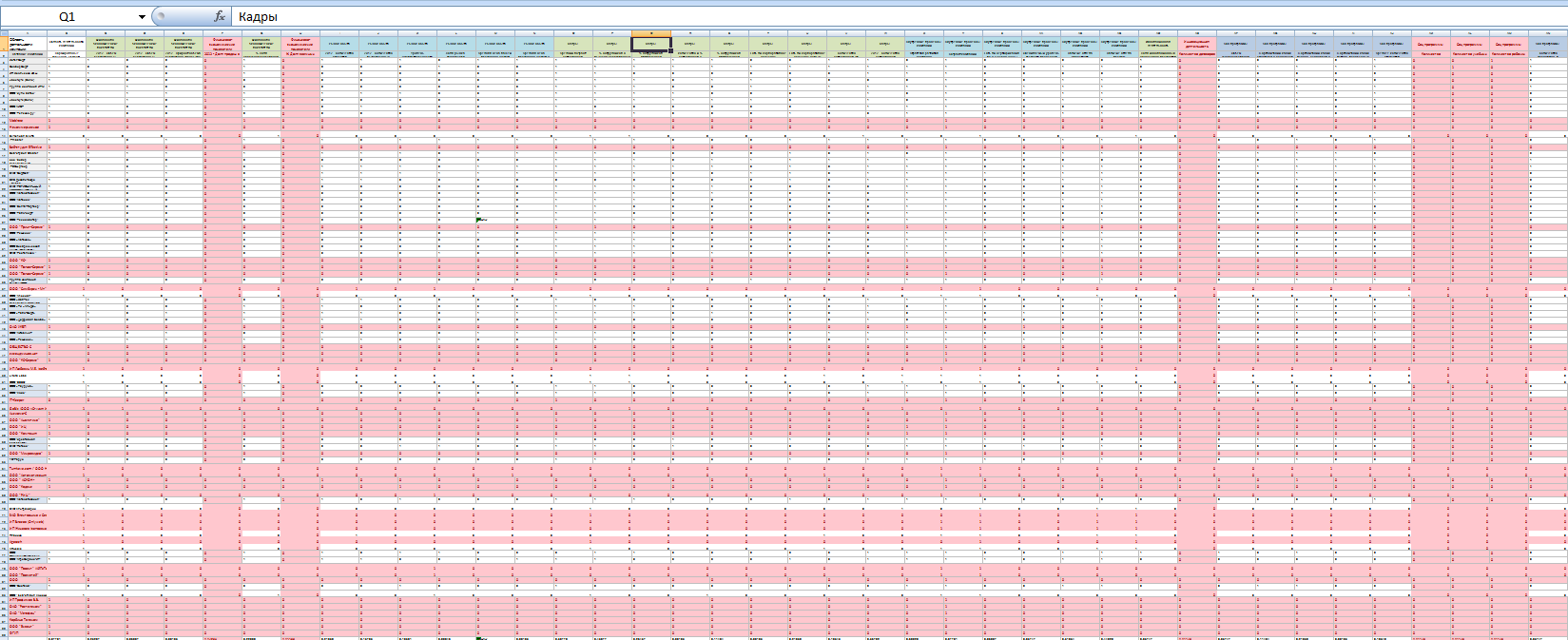


Рис. 6. Скриншот таблицы 1. Красным отмечены предприятия и данные с недостаточной плотностью заполнения (ниже выбранного уровня).

В результате редукции количество обрабатываемых предприятий сократилось с 87 до 49 , количество обрабатываемых параметров сократилось с 39 до 33. В результате мы получили матрицу данных размером и готовы приступать к кластеризации данных.

## § 2.3. Методы кластеризации, применяемые при написании программы «Кластеризация ИТ-предприятий»

В программном продукте «Кластеризация ИТ-предприятий» реализованы центроидный метод (centroid method), принадлежащий к группе дисперсионных методов иерархической кластеризации, а также неиерархический FCM-метод, позволяющий производить нечеткую кластеризацию [14].

**Реализация центроидного метода** [9], [10], [14]. В центроидном методе (методе взвешенных групп) расстояние между двумя кластерами представляет собой расстояние между их центроидами (средними для всех переменных). Каждый раз объекты группируют и вычисляют новый центроид.

Расстояние между двумя кластерами определяется как евклидово расстояние между центрами (средними) этих кластеров:

*.*

Кластеризация идет поэтапно на каждом из n–1 шагов объединяют два кластера и , имеющие минимальное значение . Если много больше  ,  то центры объединения двух кластеров близки друг к другу и характеристики  второго кластера при объединении кластеров практически игнорируются.

**Алгоритм центроидного метода.**

Шаг 1. Инициализация.

Выбираются следующие параметры: необходимое количество кластеров , ; мера расстояний как евклидово расстояние.

Изначально каждое предприятие – отдельный кластер. Список предприятий – список кластеров .

Шаг 2. Строим матрицу расстояний между центрами кластеров.

;

;

Шаг 3. Определяем минимальное расстояние

; ;

Объединим два кластера с минимальным расстоянием в один:

*=> …*

Повторяем шаги 2 и 3.

Шаг 4. В качестве критерия остановки берётся конечное число кластеров, заданное пользователем.

В программе предусмотрено графическое изображение результатов процесса  последовательной кластеризации.

С помощью дендограммы можно графически изобразить процедуру кластеризации при условии, что эта процедура оперирует только  с элементами матрицы расстояний или сходства.

В дендограмме ИТ-предприятия  располагаются вертикально слева, результаты  кластеризации – справа. Значения расстояний или сходства, отвечающие строению новых кластеров, изображаются по горизонтальной прямой поверх дендограмм.



Рис. 7. Пример дендограммы.

На рис. показан пример  дендограммы. Рис соответствует случаю шести ИТ-предприятий (n=6) и k характеристик (признаков). ИТ-предприятия А и С наиболее близки и поэтому объединяются в один кластер на уровне близости, равном 0,9. Объекты D и Е объединяются  при уровне 0,8. Теперь имеем 4 кластера: (А, С), (F), (D, E), (B).

Далее образуются кластеры (А, С, F) и (E, D, B), соответствующие уровню близости, равному 0,7 и 0,6. Окончательно все объекты группируются в один кластер при уровне 0,5.

Вид дендограммы зависит от выбора меры сходства  или расстояния между объектом  и кластером.

При решении вопроса о количестве кластеров нет твердых правил, но можно руководствоваться следующим:

–пользователь из практических соображений может захотеть получить конкретное число кластеров;

– в иерархической кластеризации в качестве критерия остановки можно использовать изначально заданное расстояние между кластерами.

**Реализация FCM-метода при кластеризации ИТ-предприятий.** Алгоритм нечеткой кластеризации называют FСМ-алгоритмом. Целью FСМ-алгоритма кластеризации является автоматическая классификация множества объектов, которые задаются векторами признаков в пространстве признаков. Такой алгоритм определяет кластеры и соответственно классифицирует объекты. Кластеры представляются нечеткими множествами, и, кроме того, границы между кластерами также являются нечеткими [14].

FСМ-алгоритм кластеризации предполагает, что объекты принадлежат всем кластерам с определенной функцией принадлежности. Степень принадлежности определяется расстоянием от объекта до соответствующих кластерных центров. Данный алгоритм итерационно вычисляет центры кластеров и новые степени принадлежности объектов.

Для заданного множества входных векторов и выделяемых кластеров предполагается, что любой принадлежит любому с принадлежностью , где – номер кластера, а – входного вектора. Принимаются во внимание следующие условия нормирования для :

 .

Цель алгоритма – минимизация суммы всех взвешенных расстояний :

,

где – фиксированный параметр, задаваемый перед итерациями.

Для достижения вышеуказанной цели необходимо решить следующую систему уравнений:

,  .

Совместно с условиями нормирования данная система дифференциальных уравнений имеет следующее решение:



Алгоритм нечеткой кластеризации выполняется по шагам.

Шаг 1. Инициализация.

Выбираются следующие параметры: необходимое количество кластеров , ; мера расстояний, как евклидово расстояние; фиксированный параметр (обычно ~1,5); начальная (на нулевой итерации) матрица принадлежности объектов с учетом заданных начальных центров кластеров .

Шаг 2. Регулирование позиций центров кластеров.

На -м итерационном шаге при известной матрице  вычисляется  в соответствии с вышеприведенным решением системы дифференциальных уравнений.

Шаг 3. Корректировка значений принадлежности

Учитывая известные , вычисляются , если в противном случае:



Шаг 4. Остановка алгоритма.

Алгоритм нечеткой кластеризации останавливается при выполнении следующего условия:

,

где  – матричная норма (например, евклидова норма); – заранее задаваемый уровень точности.

## Выводы.

Исследование ИТ-сектора – процесс сложный и неоднозначный, единой методики и показателей в аналитических материалах анализа ИТ-сектора не существует. Возникают проблемы выбора критериев кластеризации ИТ-предприятий и соответствующих им параметров.

На основании монографии Н.Г. Ярушкиной, Т.В. Афанасьевой и др. «Исследование ИТ-кластера Ульяновской области» [22] мы разделили параметры, характеризующие ИТ-предприятия на группы: область деятельности компании; финансово-экономические показатели; рынок сбыта; кадровые показатели; внутренние процессы компании; инновационная деятельность; социальные программы и образовательная деятельность.

Обрабатывались анкетные данные 87 ИТ-предприятий Ульяновской области. При анализе и обработке анкетных данных решались следующие проблемы.

– Кластеризация требует объединения близких по смысловому значению параметров в группы и сравнения объектов (ИТ-предприятий) по этим группам параметров.

– Анкетные данные разнородны. Среди них числа (рубли, количество сотрудников, проценты и т.д.), термины. Требуется перевод лингвистических переменных в числовые, требуется приведение всех данных к одному типу единиц измерения.

– Велик разброс данных, требуется их нормировка.

–Часть вопросов анкеты руководители предприятий оставили без ответа. Требуется алгоритм уменьшения размерности матрицы, исключения из рассмотрения предприятий и вопросов анкеты с нулевыми или малыми значениями параметров.

Описаны алгоритмы иерархического центроидного метода кластеризации, а также неиерархического FCM-метода, позволяющего производить нечеткую кластеризацию, которые реализованы в программном продукте «Кластеризация ИТ-предприятий».

# Глава 3. Описание программного продукта «Кластеризация ИТ-предприятий»

## § 3.1. Описание функций программного продукта «Кластеризации ИТ-предприятий»

Программный продукт предназначен для кластеризации ИТ-предприятий, для обработки экономических данных и их анализа, группировки, распознавания и представления в удобном для пользователя виде.

Производится кластеризация по следующим группам критериев: область деятельности компании; финансово-экономические показатели; рынок сбыта; кадровые показатели; внутренние процессы компании; инновационная деятельность; социальные программы и образовательная деятельность. Каждый из критериев представляет собой вектор, координатами которого являются параметры нижнего иерархического уровня, полученные как результат обработки анкеты, состоящей из 39 вопросов. Обрабатываются как числовые, так и лингвистические данные. Предусмотрена возможность выбора части критериев, возможно расширение числа критериев. В интерфейсе программного продукта используется экономическая терминология предметной области «Кластеризация IT-предприятий».

Потребителями данного программного продукта могут выступить экономисты, социологи, фирмы, лица, занимающиеся производством и продажей в сфере IT, которым полезно сегментирование рынка товаров и услуг, структурирование рынка, основанное на неоднородности и нечеткости показателей и критериев. Такой анализ необходим для дифференцированного маркетинга.

Программный продукт позволяет группировать данные иерархическим **центроидным методом кластерного анализа, неиерархическим FCM-методом,** предназначенным для нечеткой кластеризации. Выполнена программная реализация методов создания, сохранения, загрузки, обработки кластеров.

Программный продукт позволяет: визуализировать данные на экране в виде дерева, позволяющего получить результат кластеризации в виде иерархической структуры, по которой можно судить о ходе выполнения кластеризации, выводить на экран результаты анализа данных в виде таблиц, графиков и диаграмм.

Используется программный код, который может выполняться на различном HardWare, без его предварительного переноса на «родной язык» устройства. Нами был выбран C#. Код на C# будет работать везде, где установлен «C# .NET».

Сохранение и загрузка данных происходит в формате \*.SOF (Serializable Object Files).

Предусмотрена возможность использования кластерного анализа в сочетании с другими методами многомерного анализа, возможность использования различных метрик; возможность включения программного продукта в состав более сложной автоматической или автоматизированной системы для экономического анализа.

Программа может быть интегрирована в состав более сложной автоматической или автоматизированной вычислительной системы (автоматической системы управления, прогнозирования и т.д.), размеры которой могут быть как в пределах одной машины, в пределах локальной, глобальной сети, с оповещением через уже существующие средства связи (sms, Email, …).

Подобных программных продуктов для сегментации рынка IT нами не найдено.

## § 3.2. Руководство пользователю

После запуска программного продукта «Кластеризация ИТ-предприятий» появляется главная форма. На главной форме есть меню: файл, инструменты, кластеризация.

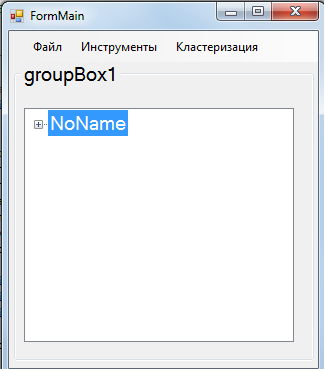
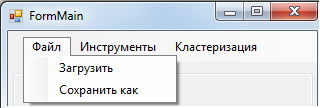


Рис. 8. Главная форма.

На рисунке 8 представлено подменю «Файл: загрузить, сохранить». На рисунках 9 и 10 представлены «Главная форма, меню «Файл»», «Главная форма/Файл/Загрузить(Открыть)».

  
Рис. 9. Главная форма, меню «Файл»

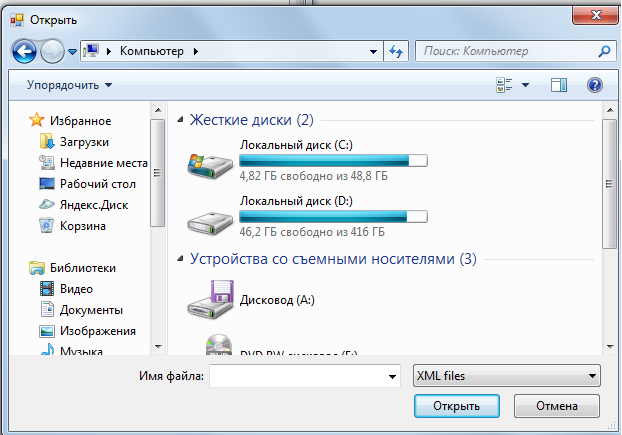


Рис. 10. Главная форма/Файл/Загрузить(Открыть)

Похожее диалоговое окно открывается и при нажатии на кнопку «Сохранить». После загрузки главная форма приобретает вид, представленный на рис. 11.

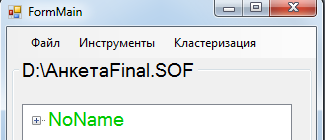


Рис.11 Главная форма после загрузки данных анкеты

Появился путь к загруженному файлу, его имя и расширение SOF (Serializable Object Files).

На главной форме реализованы три способа представления данных: дерево, меню критериев и параметров, таблица. Обсудим каждый из этих способов.

На рис.12 представлен пример дерева. Данные доступны только для просмотра. Данные представлены в виде иерархической структуры. Первый иерархический уровень ветвей дерева содержит названия ИТ-предприятий.

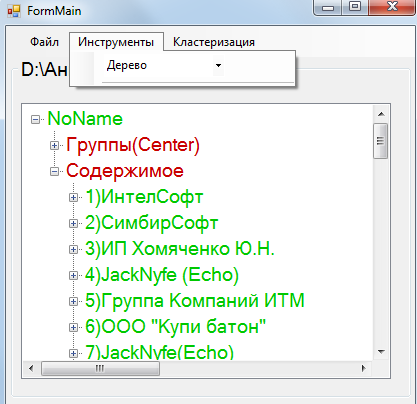
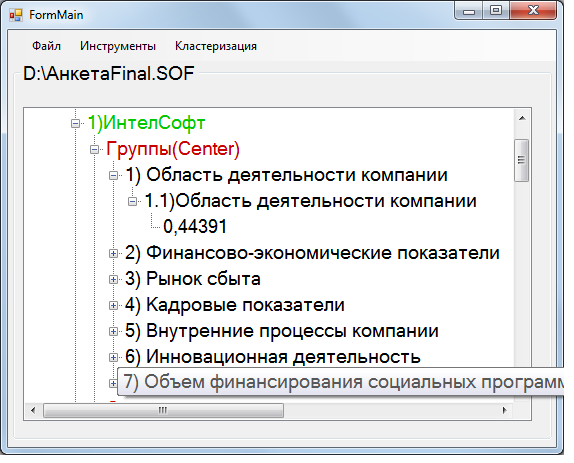


Рис. 12. Главная форма. Представление данных. Дерево.

На рис. 13 представлены ветви дерева следующего иерархического уровня – группы параметров каждого предприятия.

   
Рис. 13. Главная форма. Представление данных. Дерево – группы параметров.

Следующий иерархический уровень ветвей дерева – параметры каждой группы.

Второй способ представления данных «Меню свойств»(см. рис.14). «Меню свойств» – это элемент ASP.NET C#, позволяющий редактировать свойства экземпляра объекта.

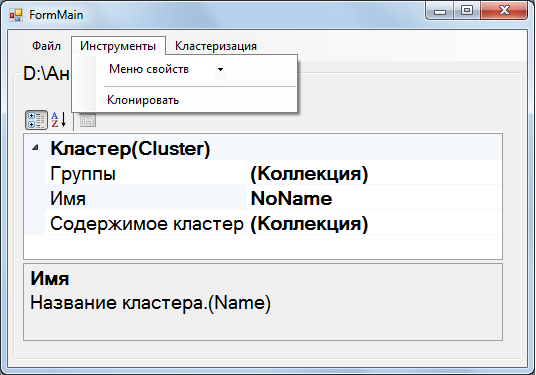


Рис. 14. Главная форма. Представление данных. Меню свойств.

В меню свойств можно редактировать данные: добавлять, удалять группы параметров, изменять имя группы, то же – для списка параметров и списка предприятий. Рекомендуется при создании нового файла задать все параметры и их группы во вкладке группы.

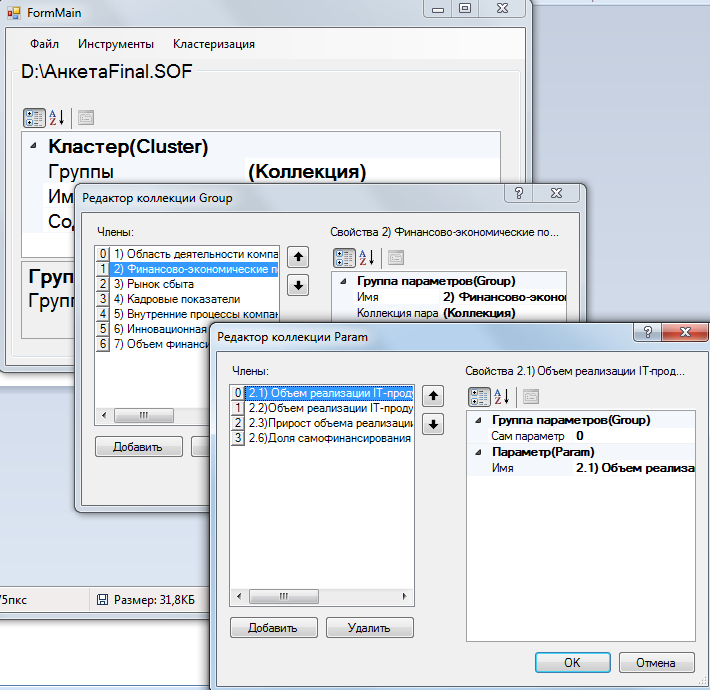
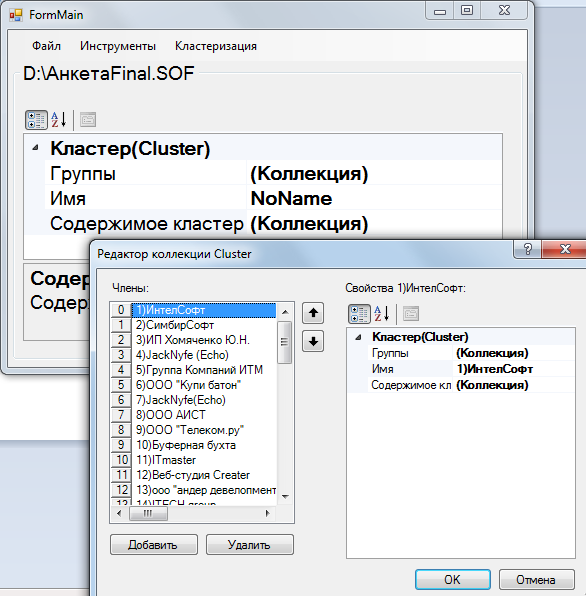


Рис. 15. Главная форма. Представление данных. Меню свойств.

Можно добавить или удалить новые предприятия в «Содержимое кластера», автоматически создать структуру для заполнения данных по каждому новому предприятию, используя инструмент «Клонировать». Это позволяет не забивать группы и их параметры для каждого предприятия, а задать их централизовано, что экономит время.

  
Рис. 16. Главная форма. Представление данных. Меню свойств.

Далее необходимо ввести все параметры. Можно это сделать при помощи инструмента «Меню свойств», но это долго и не очень удобно.  
Для этих целей более подходит инструмент «Таблица» (см. рис. 17).

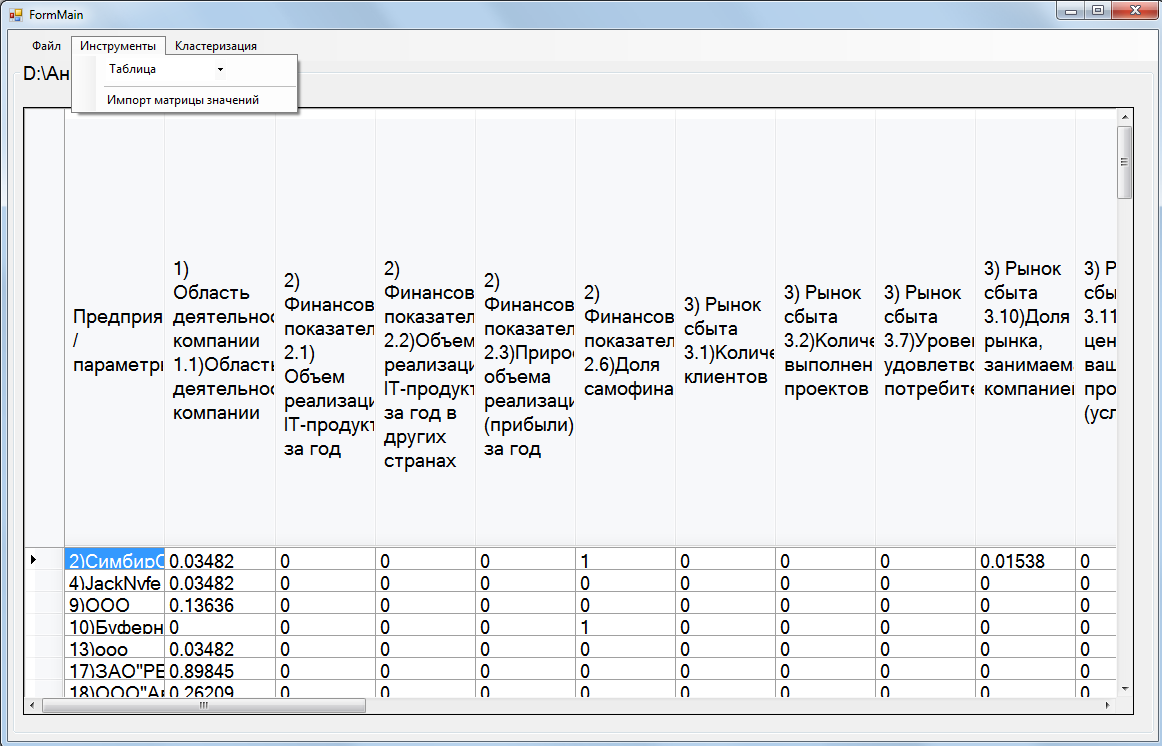


Рис. 17. Главная форма. Представление данных «Таблица».

Это представление данных имеет вид таблицы: по вертикали – список предприятий, по горизонтали – группы и соответствующие им параметры. Можно просматривать и редактировать данные в каждой отдельной ячейке (см. рис. 18). Но можно произвести «импорт таблицы» из Excel. Копируем всю таблицу, а затем просто вставляем в окно импорта данных. Нажимаем «принять».

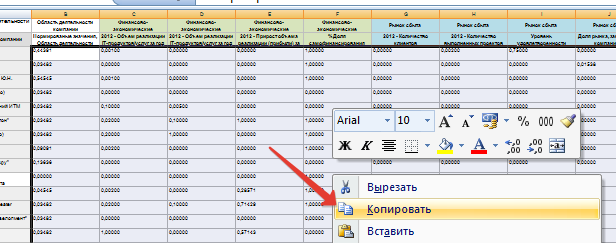


Рис.18. Главная форма. Представление данных. Таблица.

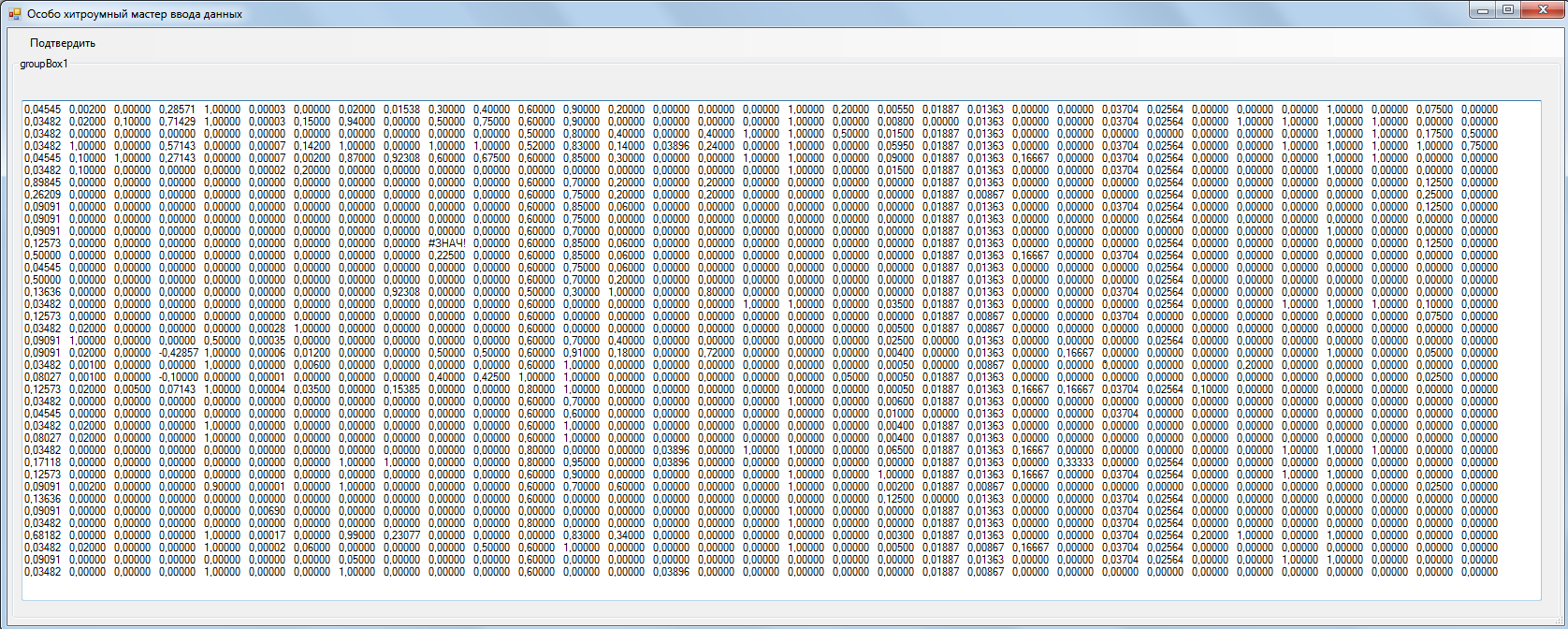
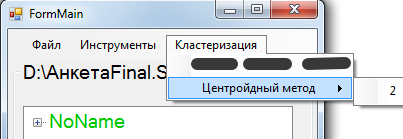


Рис. 19. Главная форма. Представление данных. Таблица.

Программа отслеживает конфликтные ситуации во время импорта, и в случае их возникновения просто не вносит изменения. Поэтому внимательно следите за размерами импортируемой таблицы, они должны соответствовать количеству предприятий и количеству параметров.

Мы рассмотрели способы просмотра, ввода и редактирования данных.  
Теперь данные можно сохранить и производить дальнейшие манипуляции.

Рассмотрим меню кластеризации.

   
Рис. 20. Главная форма. Представление данных. Таблица.

Выбираем один из представленных методов – **центроидный** (см. рис. 20). Для центроидного метода указываем количество кластеров на выходе. Результат кластеризации будет представлен на следующей форме, содержащей представления данных в виде дерева и таблицы.

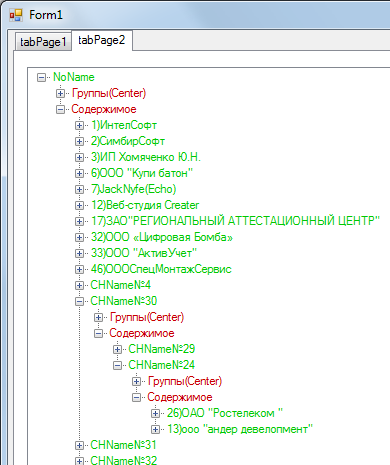
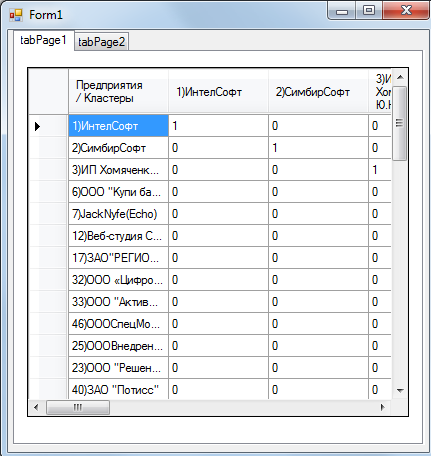


Рис. 21.Форма результатов кластеризации. Представление данных «дерево».

  
Рис. 22. Форма результатов кластеризации. Представление данных «таблица».

На рисунке 22 приведена таблица, в которой записаны результаты кластеризации (значения функции принадлежности). Эта таблица и ее анализ приводится в главе 4. По вертикали этой таблицы – предприятия, по горизонтали – кластеры. Оказалось, что некоторые предприятия образовали отдельные кластеры, что отражено в первых трех строках и столбцах таблицы приведенной на рис.22.

Для центроидного метода функция принадлежности принимает значения 0 или 1. Одно предприятие может принадлежать только к одному кластеру. Для FCM-метода значения функции принадлежности от 0 до 1.

Предусмотрена возможность выбора отдельных параметров для кластеризации из приведенного списка, кластеризация по выбранной группе параметров, по нескольким группам, по выделенному набору параметров, по всем параметрам (см. рис.23 ).

Программа позволяет использовать неиерархический **FCM-метод** (метод нечёткой кластеризации).

Этот метод использует ту же экранную форму, что и центроидный метод, что позволяет выбирать группы параметров и сами параметры для их кластеризации. Для этого метода также предусмотрена возможность ввести количество кластеров в выпадающем меню «Количество кластеров на выходе».

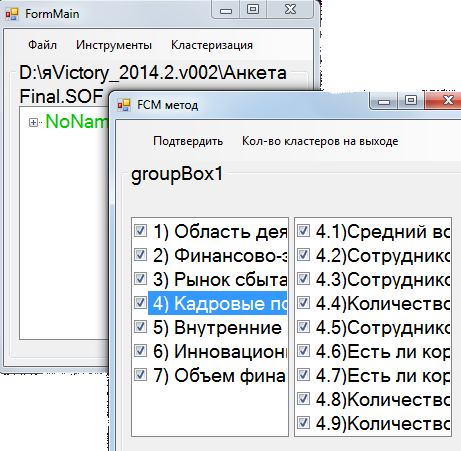
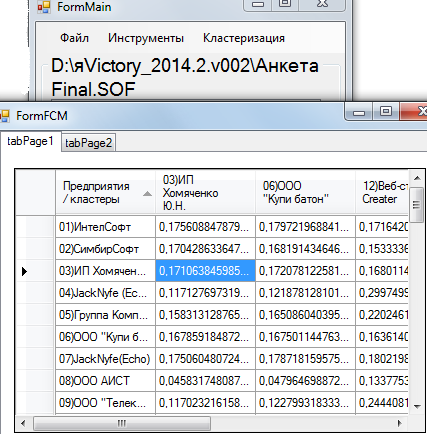


Рис. 23. Экранная форма «FCM-метод»

После подтверждения нажатием на кнопку «Подтвердить» открывается новая форма – форма результата кластеризации FCM метода.

  
Рис. 24. Экранная форма результаты кластеризации FCM-методом

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | С1 | С2 | С3 | С4 | С5 | С6 |
| 01)ИнтелСофт | 0,175608848 | 0,179722 | 0,17164204 | 0,173224743 | 0,174919771 | 0,124883 |
| 02)СимбирСофт | 0,170428634 | 0,168191 | 0,153333656 | 0,171682576 | 0,152364281 | 0,183999 |
| 03)ИП Хомяченко Ю.Н. | 0,171063846 | 0,172078 | 0,168011426 | 0,170439655 | 0,168423687 | 0,149983 |
| 04)JackNyfe (Echo) | 0,117127697 | 0,121878 | 0,299749955 | 0,114679605 | 0,269288107 | 0,077277 |
| 05)Группа Компаний ИТМ | 0,158313129 | 0,165086 | 0,220246194 | 0,154748137 | 0,201762627 | 0,099844 |
| 06)ООО "Купи батон" | 0,167859185 | 0,167501 | 0,163614086 | 0,168045244 | 0,163323205 | 0,169657 |
| 07)JackNyfe(Echo) | 0,175060481 | 0,178718 | 0,180219838 | 0,172866287 | 0,172998113 | 0,120137 |

Рис. 25. Фрагмент таблицы с результатами кластеризации FCM-методом

На рисунках 24 и 25 приведены фрагменты экранной формы с результатами кластеризации FCM-методом. На рисунках 24 и 25 приведены значения функции принадлежности предприятия кластеру. Принадлежность предприятия кластеру «нечеткая». Одной из форм контроля является суммирование элементов строки, которое должно давать 1.

§ 3.3. Руководство программисту

Структура программы «Kластеризации ИТ-предприятий». Данная программа написана на C# ASP.NET и является более поздней версий программы «Кластеризация сложных объектов». Расширение функциональных возможностей программы повлекло за собой значительные изменения используемых типов данных и архитектуры программы.

Проект программы «Kластеризации ИТ-предприятий» структурирован и содержит множество папок. Файлы и формы распределены по папкам в зависимости от функционала.

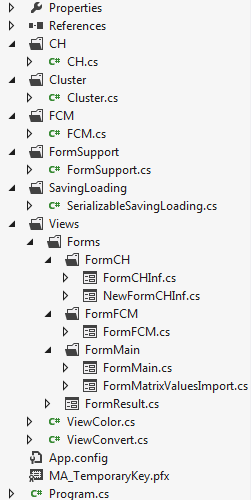


Рис. 26. Структура проекта программы

В проекте программы содержатся папки. Папки: CH, Cluster, FCM, FormSupport, SavingLoading, Views.

Первое, что увидит пользователь после загрузки программы – главная форма программы «FormMain.cs». Формы и файлы с реализацией методов визуализации данных можно найти в папке Views.

При помощи формы «FormMain.cs» пользователь может загрузить данные. Методы их загрузки сохранения можно найти в папке «SavingLoading».

Программа работает с файлами особого формата: \*.SOF (Serializable Object Files). Это связано с использованием технологии «сериализации». Данная технология позволяет записывать (считывать) в файл (поток) состояние экземпляра объекта. Такой ход позволяет абстрагировать программиста от способа хранения и передачи информации.

У данной технологии, по сравнению с предыдущей версией программы (\*.SOF против \*.XML) следующие «плюсы»:

– размер кода был сокращён с нескольких страниц до нескольких строк;  
– при дальнейшей модификации программы не возникнет необходимости модифицировать методы сохранения и загрузки данных вне зависимости от их структуры и типа (сохраняем что хотим, куда хотим и как хотим…);  
– нет возможности свободного редактирования данных вне программы. Это одновременно является плюсом и минусом: пользователю не придутся разбираться в синтаксисе \*.XML, меньше возможностей сделать ошибку при сохранении и загрузке.

Описание основного типа данных можно найти в папке «Cluster». Заложена возможность группировки параметров по группам на уровне типа данных, поддержка технологии «сериализации» (сохранение и загрузка), поддержка визуализации данных при помощи элемента формы «PropertyGrid».

Пользователь сможет увидеть данные на форме «FormMain.cs». Предоставлены 3 способа визуализации данных: при помощи элемента «TreeView», «PropertyGrid», «DataGridView». Визуализация при помощи элементов «TreeView» и «DataGridView» реализована в папке «Views» в файлах «ViewColor.cs» (здесь описание наиболее распространенных цветов) и «ViewConvert.cs» (здесь описание группы методов, осуществляющих преобразования типов данных Cluster=>TreeNode и Cluster=>DataTable). Визуализация при помощи элемента «PropertyGrid» реализована в папке «Cluster» в файле «Cluster.cs».

При необходимости пользователь сможет изменить размеры любой формы, при этом изменятся размеры элементов формы.

Папка «FormSupport» содержит «FormSupport.cs» файл реализации изменяемой геометрии форм и её элементов. Это удобно. Нет необходимости прописывать для каждой формы правила изменения геометрии. Достаточно обратиться к содержимому этой папки.

Пользователь может также изменять данные на главной форме программы («FormMain.cs»). Взаимодействие, обновление данных и их визуализация реализованы в коде этой же формы.

Пользователь может на главной форме запустить ветку программы кластеризации данных. Сценарий запуска форм и передачи между ними данных также реализован на главной форме.

В ветке программы кластеризации данных запускается форма выбора групп параметров и их параметров «NewFormCHInf.cs». В этой форме реализовано «вырезание лишней информации», а именно невыделенных параметров, а затем групп параметров. Дальше передаются данные, в которых нет и намёка на вырезанные данные. Далее, в зависимости от метода, открывается форма результатов кластеризации, в которую добавлены кластеризованные данные ещё на этапе инициализации формы. То есть форма результатов кластеризации данных содержит в себе только «доводку» до визуального представления данных.

Реализацию методов кластеризации данных можно найти в папках «CH», «FCM» – центроидный и FCM-методы, соответственно.

Это было краткое описание архитектуры программы. Более детальное описание возложено в комментариях в коде программы (см. приложение 2).

Выводы

Программный продукт предназначен для кластеризации ИТ-предприятий, для обработки экономических данных и их анализа, группировки, распознавания и представления в удобном для пользователя виде.

Производится кластеризация по следующим группам критериев: область деятельности компании; финансово-экономические показатели; рынок сбыта; кадровые показатели; внутренние процессы компании; инновационная деятельность; социальные программы и образовательная деятельность. Обрабатываются как числовые, так и лингвистические данные. Предусмотрена возможность выбора критериев, возможно расширение их числа. В интерфейсе программного продукта используется экономическая терминология предметной области «Кластеризация IT-предприятий».

Программный продукт позволяет группировать данные иерархическим центроидным методом кластерного анализа, неиерархическим FCM-методом, предназначенным для нечеткой кластеризации. Выполнена программная реализация методов создания, сохранения, загрузки, обработки кластеров.

Программный продукт позволяет: визуализировать данные на экране в виде дерева, позволяющего получить результат кластеризации в виде иерархической структуры, по которой можно судить о ходе выполнения кластеризации, выводить на экран результаты анализа данных в виде таблиц, графиков и диаграмм.

Используется программный код, который может выполняться на различном HardWare, без его предварительного переноса на «родной язык» устройства. Нами был выбран C#. Код на C# будет работать везде, где установлен «C# .NET». Сохранение и загрузка данных происходит в формате \*.SOF (Serializable Object Files).

Потребителями программного продукта могут выступить экономисты, социологи, фирмы, лица, занимающиеся производством и продажей в сфере IT. Кластерный анализ необходим для дифференцированного маркетинга.

# Глава 4. Анализ результатов кластеризации ИТ-предприятий Ульяновской области

Нами обрабатывались данные по 87 предприятиям Ульяновской области. Анкетирование проводилось в 2013г. Обрабатываемая анкета включала 39 вопросов. После предварительной обработки таблица данных приобрела размерность .

Предлагаемый программный продукт позволяет произвести кластеризацию по выбранным критериям, по каждой группе критериев отдельно, по нескольким группам, по всем критериям одновременно.

Рассмотрим результаты кластеризации по каждой группе критериев. Производилась кластеризация на 2, 4, 6, 8, 10, 15 кластеров.

## § 4.1. Результаты кластеризации по критерию «Область деятельности компании»

Таблица 4 представляет собой матрицу принадлежности ИТ-предприятий кластерам. Представляем результаты кластеризация на 6, 10, 15 кластеров по группе параметров «Область деятельности компании». 1 означает принадлежность ИТ-предприятия кластеру, 0 – непринадлежность. По вертикали в таблице перечислены ИТ-предприятия, по горизонтали – условные названия кластеров. По таблице 4 можно проследить, как при уменьшении числа кластеров ИТ-предприятия объединяются в более крупные кластеры.

Проанализируем матрицу для 6 кластеров таблицы 4. Самым многочисленным является кластер С5. В него вошли 88% предприятий. Типичная область деятельности этих предприятий – заказная разработка программного обеспечения: интернет-сайты, веб-приложения, разработка и сопровождение программных продуктов, интернет-провайдеры, ремонт и обслуживание техники, продажа, системная интеграция и администрирование.

Оставшиеся 12 процентов предприятий представляют редкие среди опрошенных предприятий виды деятельности:

* кластер С1 – «ИнтелСофт» – кроме разработки программных продуктов занимаются обучением ИТ-технологиям;
* кластер С2 – ИП «Хомяченко» – научные исследования, «Умный дом»;
* кластер С3 – ЗАО «Региональный аттестационный центр» – кроме разработки программных продуктов, ремонта и обслуживания техники занимается защитой информации;
* кластер С4 – ООО «Спец Монтаж Сервис» – научные исследования, телекоммуникации;
* кластер С6 – ООО «Аист» – продажа компьютерной техники, ремонт и обслуживание, системная интеграция (ERP, CRM, консультационные системы и т.д.), системное администрирование, видеонаблюдение, ЗАО «Потисс» – проектирование объектов связи.

Это позволяет заключить, что в Ульяновской области мало ИТ-предприятий занимаются обучением, научными исследованиями в ИТ-сфере и защитой информации.

Разделение предприятий на 10 кластеров даёт более тонкую дифференциацию по сфере деятельности. Например, в кластер С10 оказались выделены ИТ-предприятий, занимающиеся в основном продажей, ремонтом и обслуживанием техники. В кластер С8 попали «Ростелеком», «Ультрамарин», занимающиеся интернет-провайдом и услугами связи и т.д.

Так как в анкетах упоминалось 7 основных видов деятельности (разработка, продажа, ремонт и обслуживание, интернет и телекоммуникации, обучение, научные исследования, защита информации), и это было положено в основу кластеризации, то кластеризация на большее число кластеров (см., например, 15 кластеров) даёт дифференциацию по количеству различных видов деятельности на ИТ-предприятиях. Такое разбиение не всегда информативно.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Область деятельности компании | С1 | С2 | С3 | С4 | С5 | С6 |  | С1 | С2 | С3 | С4 | С5 | С6 | С7 | С8 | С9 | С10 |  | С1 | С2 | С3 | С4 | С5 | С6 | С7 | С8  Таблица 4 . | С9 | С10 | С11 | С12 | С13 | С14 | С15 |
| 01)ИнтелСофт | 1 | 0 | 0 | 0 | 0 | 0 |  | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |  | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 02)СимбирСофт | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 03)ИП Хомяченко | 0 | 1 | 0 | 0 | 0 | 0 |  | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |  | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 04)JackNyfe | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 05)Гр.Комп.ИТМ | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 06) "Купи батон" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 07)JackNyfe | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 08)ООО АИСТ | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |  | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 09)ООО "Телеком.ру" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |  | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10)Буферная бухта | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 11)ITmaster | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 12)Веб-студия Creater | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 13)ооо"андер девелопмент" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 14)ITECH.group | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 15)ЗАО "Софткей" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 16)B2B Дизайн «Зебра» | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 17)ЗАОРЕГ.АТТЕСТ.ЦЕНТР | 0 | 0 | 1 | 0 | 0 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 18)ООО"Автоматизация" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 19)ООО"Автоном" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 20)ООО"Волга-Партнер" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 21)ООО "Тайм-софт" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 22)ООО"Техномастер" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 23)ООО "Решение" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 24)ООО «Автоком» | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 25)ООО "ПаритетЪ" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 26)ОАО "Ростелеком " | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 27)Группа компаний Симтек | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 28)ООО "Арсенал" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 29)ОООХайТэкДевелопмент | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 30)ООО «ТК «Альфа» | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 31)ООО «Тайм-Софт» | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 32)ООО«Цифровая Бомба» | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 33)ООО "АктивУчет" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 34)ООО «Решение» | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 35)Litota Labs | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 36)ООО "ЗЕВС" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 37)ООО «Тауруна» | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 38)ООО "УмКо" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 39)ОООКреатив.разработка | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 40)ЗАО "Потисс" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 41)1С-Рарус | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 42)ООО"Автоматизация" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 43)ЗАО Ультрамарин | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44)Планикс | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 45)Artwedis | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 46)ОООСпецМонтажСервис | 0 | 0 | 0 | 1 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 47)ООО"АгроСервис-ИТ" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 48)ООО "Система" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 49)ООО " веб. Умножение" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

**Центры кластеризации**

Рис. 27. Нормы центров кластеров

при кластеризации по критерию «Область деятельности компании», .

На рисунке 27 показаны длины векторов – центров кластеризации. Нормировка векторов была такой, что часто встречающиеся виды деятельности получили наименьший вес. Поэтому при кластеризации на шесть кластеров наиболее многочисленный кластер С5 с типичными видами деятельности на первом графике имеет наименьшую длину вектора. Максимум имеет кластер С3, ИТ-предприятия которого помимо обычных видов деятельности занимаются защитой информации. ИТ-предприятие, попавшее в кластер С1, из общего кластера С5 выделилось тем, что занимается еще и обучением, что увеличило длину соответствующего вектора. На втором графике лидеры по сферам деятельности – защита информации и наука.

## § 4.2. Кластеризация по финансово-экономическим показателям

Проанализируем результаты кластеризации предприятий на 4, 6, 10 предприятий по группе параметров «Финансово-экономические показатели» (см. таблицу 5). Они показались нам наиболее информативными при анализе результатов вычислительного эксперимента.

Таблица 5

Кластеризация по критерию «Финансово-экономические показатели»

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Финансово-экономические показатели | С1 | С2 | С3 | С4 |  | С1 | С2 | С3 | С4 | С5 | С6 |  | С1 | С2 | С3 | С4 | С5 | С6 | С7 | С8 | С9 | С10 |
| 01)ИнтелСофт | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 1 |  | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 02)СимбирСофт | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 03)ИП Хомяченко Ю.Н. | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 04)JackNyfe (Echo) | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 05)Группа Компаний ИТМ | 1 | 0 | 0 | 0 |  | 1 | 0 | 0 | 0 | 0 | 0 |  | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 06)ООО "Купи батон" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 07)JackNyfe(Echo) | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 | 0 | 0 |  | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 08)ООО АИСТ | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 09)ООО "Телеком.ру" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 10)Буферная бухта | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 11)ITmaster | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 12)Веб-студия Creater | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 13)ооо"андер девелопмент" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 14)ITECH.group | 0 | 0 | 1 | 0 |  | 0 | 0 | 1 | 0 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 15)ЗАО "Софткей" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 1 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 16)B2B Дизайн-бюро«Зебра» | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 1 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 17)ЗАО"РЕГИОН.АТТЕСТ.ЦЕНТР" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 18)ООО"Автоматизация" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 19)ООО"Автоном" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 20)ООО"Волга-Партнер" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 21)ООО "Тайм-софт" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 22)ООО"Техномастер" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 23)ООО "Решение" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 24)ООО «Автоком» | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 25)ОООВнедренческий центр "ПаритетЪ" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 26)ОАО "Ростелеком " | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 27)Группа компаний Симтек | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 28)ООО "Арсенал" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 29)ООО «ХайТэк Девелопмент Групп» | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 30)ООО «ТК «Альфа» | 0 | 0 | 1 | 0 |  | 0 | 0 | 1 | 0 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 31)ООО «Тайм-Софт» | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 32)ООО «Цифровая Бомба» | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 33)ООО "АктивУчет" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 34)ООО «Решение» | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 35)Litota Labs | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 36)ООО "ЗЕВС" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 37)ООО «Тауруна» | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 38)ООО "УмКо" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 39)ООО "Креативная разработка" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 40)ЗАО "Потисс" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 41)1С-Рарус | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 42)ООО"Автоматизация" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 43)ЗАО Ультрамарин | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 44)Планикс | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 45)Artwedis | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 46)ОООСпецМонтажСервис | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 47)ООО"АгроСервис-ИТ" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 48)ООО "Система" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 49)ООО " веб-студия Умножение" | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |

При кластеризации на 4 кластера самый многочисленный кластер – C4. ИТ-предприятиям этого кластера свойственны высокая доля самофинансирования и прирост объема реализации (прибыли) за год, но не самый высокий объем реализации ИТ-продуктов/услуг. ИТ-предприятия этого кластера можно охарактеризовать, как быстро растущие и динамично развивающиеся предприятия.

Кластер С3 и С2 – наиболее прогрессирующие по объему реализации и по приросту объема реализации (прибыли) за год предприятия.

В кластер С3 вошли ITECH.group, ООО «ТК «Альфа». Сильнейшие стороны этого кластера: объем реализации IT-продуктов/услуг за год, прирост объема реализации (прибыли) за год. В кластер С2 входит JackNyfe(Echo). Для этого кластера высоки показатели: объем реализации IT-продуктов/услуг за год в других странах и доля самофинансирования.

Таблица 6

Нормированные значения параметров группы «Финансово-экономические показатели» при кластеризации на 4 и 6 кластеров

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Кластеры  Показатели | С1 | С2 | С3 | С4 |  | С1 | С2 | С3 | С4 | С5 | С6 |
| 2.1) Объем реализации IT-продуктов/услуг за год | 0,1 | 0,2 | 1 | 0,007 |  | 0,1 | 0,2 | 1 | 0,1 | 0,0002 | 0,017 |
| 2.2) Объем реализации IT-продуктов/услуг за год в других странах | 0,005 | 1 | 0 | 0,022 |  | 0,005 | 1 | 0 | 0,5 | 0 | 0,022 |
| 2.3)Прирост объема реализации (прибыли) за год | 0,429 | 0,429 | 0,714 | 0,458 |  | 0,429 | 0,429 | 0,714 | 0,564 | 0,421 | 0,575 |
| 2.6)Доля самофинансирования | 0 | 1 | 0,25 | 0,375 |  | 0 | 1 | 0,25 | 0 | 0,227 | 0,892 |

В столбцах таблицы 6 приведены координаты векторов – центров кластеров при кластеризации на 4 и 6 кластеров по группе «Финансово-экономические показатели». Нормировка векторов-столбцов произведена так, что каждая координата обозначает долю от предприятия, имеющего максимальное значение соответствующего параметра. Для центра кластера берется средняя по кластеру величина показателя.

По таблице 6 видно, что кластер С2 (JackNyfe(Echo) имеет максимальное среди всех ИТ-предприятий Ульяновской области объем реализации IT-продуктов/услуг за год в других странах (50-100 млн. руб) и долю самофинансирования 100%.

Кластер С3 (ITECH.group, ТК «Альфа») лидер по объему реализации IT-продуктов/услуг за год (100-500 млн.руб).

Кластер С1 (группа Компаний ИТМ) отличается нулевым самофинансированием (или оно не указано в анкете), невысоким объемом реализации IT-продуктов/услуг за год (10-50 млн.руб – 10% по сравнению с лидером) и объем реализации IT-продуктов/услуг за год (0-500 тыс.руб в других странах), средним по области приростом объема реализации (прибыли) за год.

Самый многочисленный кластер С4 (90% ИТ-предприятий) имеет средние значения показателей: объем реализации IT-продуктов/услуг за год в промежутке от 0,7 до 3,5 млн.руб; объем реализации IT-продуктов/услуг за год в других странах составляет часть этого промежутка 1,1-2,2 млн.руб; прирост объема реализации (прибыли) за год – 46 %; доля самофинансирования – 37,5 %.

При кластеризации на шесть кластеров кластеры С1, С2, С3 остались те же. Кластер С4 разделился на три. Среди них С4 (ЗАО «Софткей», B2B Дизайн-бюро «Зебра») отличается более высоким объемом реализации в других странах (показатель 0,5 выше среднего по бывшему кластеру С4, где он был 0,022), доля самофинансирования – 0%.

С6 (ИнтелСофт, ООО «Купи батон», Веб-студия Creater, ООО «ХайТэк Девелопмент Групп», ООО «Тайм-Софт», ООО «Решение», ООО «Тауруна», ООО «УмКо», ООО «АгроСервис-ИТ») выделяет из бывшего кластера то, что здесь в десять раз выше объем реализации и высокая доля самофинансирования (в среднем – 89%).

Большинство ИТ-предприятий остались в кластере С5, где средний объем реализации продукции, указанный в анкете (20-100 тыс. руб, при среднем показателе по бывшему кластеру – 700 тыс. -3,5 млн. руб) и нулевой объем реализации в других странах. Показатели в этой группе могут оказаться заниженными, так как часть предприятий этого кластера, видимо из соображений коммерческой тайны, не заполнило соответствующую графу анкеты.

**Центры кластеризации**

**Центры кластеризации**

Рис. 28. Нормы центров кластеров

при кластеризации по критерию «Финансово-экономические показатели».

На рисунке 28 представлены длины векторов – центров кластеров при кластеризации по критерию «Финансово-экономические показатели» при . Первый график показывает, что лидером по финансово-экономическим показателям при является кластер С2 (JackNyfe(Echo), аутсайдер С1 (группа Компаний ИТМ, судя по заполненным или незаполненным графам анкеты).

При на первом месте по финансово-экономическим показателям – кластер С3 (JackNyfe(Echo), на втором – С8 (ITECH.group, ООО «ТК «Альфа»), на третьем – С1 (ИнтелСофт, ООО «Решение»), С6 (ООО «Тауруна», ООО «УмКо», ООО»АгроСервис-ИТ»), С10 (ООО «Купи батон», Веб-студия Creater, ООО «ХайТэк Девелопмент Групп», ООО «Тайм-Софт»).

Рис. 29. Доля ИТ-предприятий, входящих в кластеры

при кластеризации на 4,6, 10 кластеров по финансово-экономическим показателям

На рисунке представлена доля ИТ-предприятий, входящих в кластеры при кластеризации на 4,6, 10 кластеров.

## § 4.3 Кластеризация по критерию «Рынок сбыта»

Проанализируем результаты кластеризации предприятий на 6, 8, 10 кластеров по группе параметров «Рынок сбыта». Они показались нам наиболее информативными при анализе результатов вычислительного эксперимента. В таблице 7 приведены матрицы принадлежности ИТ-предприятий кластерам при .

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Рынок сбыта | С1 | С2 | С3 | С4 | С5 | С6 |  | С1 | С2 | С3 | С4 | С5 | С6 | С7 | С8 |  | С1 | С2 | С3 | С4 | С5 | С6  Таблица 7 | С7 | С8 | С9 | С10 |
| 01)ИнтелСофт | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 02)СимбирСофт | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 03)ИП Хомяченко Ю.Н. | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 04)JackNyfe (Echo) | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 05)Группа Компаний ИТМ | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 06)ООО "Купи батон" | 1 | 0 | 0 | 0 | 0 | 0 |  | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |  | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 07)JackNyfe(Echo) | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 08)ООО АИСТ | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 09)ООО "Телеком.ру" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 10)Буферная бухта | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 11)ITmaster | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 12)Веб-студия Creater | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 13)ооо "андер девелопмент" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 14)ITECH.group | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 15)ЗАО "Софткей" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 16)B2B Дизайн-бюро «Зебра» | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 17)ЗАО"РЕГИОНАЛЬНЫЙ АТТЕСТАЦИОННЫЙ ЦЕНТР" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 18)ООО"Автоматизация" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 19)ООО"Автоном" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 20)ООО"Волга-Партнер" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 21)ООО "Тайм-софт" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 22)ООО"Техномастер" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 23)ООО "Решение" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 24)ООО «Автоком» | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 25)ОООВнедренческий центр "ПаритетЪ" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 26)ОАО "Ростелеком " | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 27)Группа компаний | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 28)ООО "Арсенал" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 29)ООО «ХайТэк Девелопмент Групп» | 0 | 1 | 0 | 0 | 0 | 0 |  | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |  | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 30)ООО «ТК «Альфа» | 0 | 0 | 1 | 0 | 0 | 0 |  | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |  | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 31)ООО «Тайм-Софт» | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |  | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 32)ООО «Цифровая Бомба» | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 33)ООО "АктивУчет" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 34)ООО «Решение» | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 35)Litota Labs | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 36)ООО "ЗЕВС" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 37)ООО «Тауруна» | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 38)ООО "УмКо" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 39)ООО "Креативная разработка" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 40)ЗАО "Потисс" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 41)1С-Рарус | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 42)ООО"Автоматизация" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 43)ЗАО Ультрамарин | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 44)Планикс | 0 | 0 | 0 | 1 | 0 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 45)Artwedis | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 46)ОООСпецМонтажСервис | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 47)ООО"АгроСервис-ИТ" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 48)ООО "Система" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 49)ООО " веб-студия Умножение" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |

При кластеризации на 6 кластеров самый многочисленный кластер – C6. ИТ-предприятиям этого кластера свойственны высокая плотность заполнения анкеты в данной группе параметров. Для этого кластера характерны наибольшие значения по параметрам «Средняя цена вашей продукции (услуг)» и «Средняя цена продукции (услуг) у ваших конкурентов рублей / час», но наименьшее значение по параметру «Количество клиентов».

Прочие кластеры представлены каждый – одним ИТ-предприятием.

В числе лидеров:

– кластер С1 (ООО «Купи батон»), обладающий наибольшим количеством клиентов, при одном разработанном проекте они его продали 2898365 клиентам;

–кластер С2 (ООО «ХайТэк Девелопмент Групп»), обладающий наибольшим значением по параметру «Количество выполненных проектов», они выполнили 1000 проектов для 800 клиентов;

– кластер С5 (ООО «СпецМонтажСервис»), имеющий 500 клиентов.

У кластеров С3 (ООО «ТК «Альфа»), С4 («Планикс») единственным заполненным параметром оказалось «Количеством клиентов».

Центры кластеризации

Рис. 30 . Кластеризация по критерию «Рынок сбыта» (нормы центров кластеров,

доля ИТ-предприятий, входящих в кластер),

## § 4.4. Кадровые показатели.

Рассмотрим кластеризацию предприятий по группе кадровых показателей. В таблице 8 представлены матрицы принадлежности ИТ-предприятий кластерам .

При кластеризации на 6 кластеров – кластер C1 состоит из одного ИТ-предприятия – ИнтелСофт. В нем при 4 сотрудниках максимальная доля сотрудников с высшем образованием (75%), с научными степенями (25%), выпускников последних лет и студентов (25%).

Центры кластеров

Рис. 31. Кластеризация по критерию «Инновационная деятельность» (нормы центров кластеров, доля ИТ-предприятий, входящих в кластер,

В отдельные кластеры выделились:

– кластер C3 - ITECH.group 119 сотрудников, 83% из которых с высшим образованием, средний возраст типичен для данной отрасли – 26 лет, имеется корпоративное обучение новым технологиям.

–кластер C2 - ИП Хомяченко Ю.Н. 1 сотрудник 30-40 лет с высшим образованием.

– кластер C4 - ООО «АктивУчет» 1 сотрудник 40-50 лет с высшим образованием.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Кадровые показатели | С1 | С2 | С3 | С4 | С5 | С6 |  | С1 | С2 | С3 | С4 | С5 | С6 | С7 | С8 |  | С1 | С2 | С3 | С4  Таблица 8 | С5 | С6 | С7 | С8 | С9 | С10 |
| 01)ИнтелСофт | 1 | 0 | 0 | 0 | 0 | 0 |  | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |  | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 02)СимбирСофт | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 03)ИП Хомяченко Ю.Н. | 0 | 1 | 0 | 0 | 0 | 0 |  | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |  | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 04)JackNyfe (Echo) | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 05)Группа Компаний ИТМ | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 06)ООО "Купи батон" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 07)JackNyfe(Echo) | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 08)ООО АИСТ | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 09)ООО "Телеком.ру" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 10)Буферная бухта | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 11)ITmaster | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 12)Веб-студия Creater | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 13)ооо "андер девелопмент" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 14)ITECH.group | 0 | 0 | 1 | 0 | 0 | 0 |  | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |  | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15)ЗАО "Софткей" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 16)B2B Дизайн-бюро «Зебра» | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 17)ЗАО"РЕГИОНАЛЬНЫЙ АТТЕСТАЦИОННЫЙ ЦЕНТР" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 18)ООО"Автоматизация" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 19)ООО"Автоном" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 20)ООО"Волга-Партнер" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 21)ООО "Тайм-софт" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 22)ООО"Техномастер" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 23)ООО "Решение" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 24)ООО «Автоком» | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 25)ОООВнедренческий центр "ПаритетЪ" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 26)ОАО "Ростелеком " | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 27)Группа компаний Симтек | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 28)ООО "Арсенал" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 29)ООО «ХайТэк Девелопмент Групп» | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 30)ООО «ТК «Альфа» | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 31)ООО «Тайм-Софт» | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 32)ООО «Цифровая Бомба» | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 33)ООО "АктивУчет" | 0 | 0 | 0 | 1 | 0 | 0 |  | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |  | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 34)ООО «Решение» | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 35)Litota Labs | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 36)ООО "ЗЕВС" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 37)ООО «Тауруна» | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 38)ООО "УмКо" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 39)ООО "Креативная разработка" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 40)ЗАО "Потисс" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 41)1С-Рарус | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 42)ООО"Автоматизация" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 43)ЗАО Ультрамарин | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 44)Планикс | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 45)Artwedis | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 46)ОООСпецМонтажСервис | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 47)ООО"АгроСервис-ИТ" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 48)ООО "Система" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 49)ООО " веб-студия Умножение" | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |

Таблица 9

Нормированные значения параметров группы «Кадровые показатели» при кластеризации на 6 кластеров

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | С1 | С2 | С3 | С4 | С5 | С6 |
| 4.1)Средний возраст сотрудников | 0,6 | 0,8 | 0,52 | 1 | 0,555882 | 0,663889 |
| 4.2)Сотрудников с выс. Образ. | 0,75 | 1 | 0,83 | 1 | 0,62733 | 0,655208 |
| 4.3)Сотрудников – студентов | 0,5 | 0 | 0,14 | 0 | 0,089237 | 0,194444 |
| 4.4)Количество и сотрудников с науч. степенями | 0,32468 | 0 | 0,03896 | 0 | 0,002256 | 0,111653 |
| 4.5)Сотрудников – выпускников 2012 г. | 1 | 0 | 0,24 | 0 | 0,035901 | 0,128528 |
| 4.6)Корпоративное обучение ин. языкам | 0 | 0 | 0 | 0 | 0,086857 | 0,493056 |
| 4.7)Корпоративное обучение новым технологиям | 0 | 0 | 1 | 0 | 0,308364 | 0,5 |
| 4.8)Сотрудники (за последние 5лет), участвующие в конференциях… | 0,1 | 0 | 0 | 0,05 | 0,021714 | 0,204167 |
| 4.9)Количество сотрудников | 0,002 | 0,0005 | 0,0595 | 0,0005 | 0,039363 | 0,013531 |

В кластере C6 (СимбирСофт, JackNyfe (Echo), Группа Компаний ИТМ, ООО «Купи батон», ООО «Андер девелопмент», ОАО «Ростелеком», ООО «Решение», ООО «Креативная разработка», ЗАО «Потисс», Artwedis) от 10 до 70 сотрудников возраста 25-30 лет, из которых примерно 80 % с высшим образованием. В половине компаний этого кластера имеется корпоративное обучение.

В кластере C5 сосредоточена основная масса предприятий (72 %). У данного кластера низкий интегративный показатель, характеризующий состояние кадров (см. рис.31 и таблицу 9).

§ 4.5. Внутренние процессы компании. Инновационная деятельность. Социальные программы и образовательная деятельность.

По группам критериев «внутренние процессы компании», «инновационная деятельность», «социальные программы и образовательная деятельность» наиболее информативной оказалась кластеризация на 4-6 кластеров.

Таблица 10

Кластеризация по группам критериев «внутренние процессы компании», «инновационная деятельность», «социальные программы и образовательная деятельность»

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Внутренние процессы компании | | | | Инновационная деятельность | | | | |  | Социальные программы | | | |
|  | С1 | С2 | С3 | С4 |  | С1 | С2 | С3 | С4 |  | С1 | С2 | С3 | С4 |
| 01)ИнтелСофт | 1 | 0 | 0 | 0 |  | 0 | 0 | 0 | 1 |  | 0 | 1 | 0 | 0 |
| 02)СимбирСофт | 0 | 0 | 0 | 1 |  | 0 | 1 | 0 | 0 |  | 1 | 0 | 0 | 0 |
| 03)ИП Хомяченко Ю.Н. | 0 | 0 | 0 | 1 |  | 0 | 0 | 1 | 0 |  | 0 | 1 | 0 | 0 |
| 04)JackNyfe (Echo) | 1 | 0 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 05)Группа Компаний ИТМ | 0 | 0 | 1 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 06)ООО "Купи батон" | 1 | 0 | 0 | 0 |  | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 1 |
| 07)JackNyfe(Echo) | 1 | 0 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 08)ООО АИСТ | 0 | 0 | 1 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 09)ООО "Телеком.ру" | 0 | 0 | 1 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 10)Буферная бухта | 1 | 0 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 11)ITmaster | 0 | 0 | 1 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 12)Веб-студия Creater | 1 | 0 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 0 | 1 | 0 |
| 13)ооо "андер девелопмент" | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 14)ITECH.group | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 15)ЗАО "Софткей" | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 16)B2B Дизайн-бюро «Зебра» | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 17)ЗАО"РЕГИОН. АТТЕСТ. ЦЕНТР" | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 18)ООО"Автоматизация" | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 19)ООО"Автоном" | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 20)ООО"Волга-Партнер" | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 21)ООО "Тайм-софт" | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 22)ООО"Техномастер" | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 23)ООО "Решение" | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 24)ООО «Автоком» | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 25)ООО центр "ПаритетЪ" | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 26)ОАО "Ростелеком " | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 27)Группа компаний Симтек | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 28)ООО "Арсенал" | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 29)ООО «ХайТэк Девелопмент Групп» | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 30)ООО «ТК «Альфа» | 1 | 0 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 31)ООО «Тайм-Софт» | 1 | 0 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 32)ООО «Цифровая Бомба» | 1 | 0 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 0 | 0 | 1 |
| 33)ООО "АктивУчет" | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 34)ООО «Решение» | 0 | 1 | 0 | 0 |  | 1 | 0 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 35)Litota Labs | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 36)ООО "ЗЕВС" | 1 | 0 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 37)ООО «Тауруна» | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 38)ООО "УмКо" | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 39)ООО "Креативная разработка" | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 40)ЗАО "Потисс" | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 41)1С-Рарус | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 42)ООО"Автоматизация" | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 43)ЗАО Ультрамарин | 1 | 0 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 44)Планикс | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 45)Artwedis | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 46)ОООСпецМонтажСервис | 0 | 1 | 0 | 0 |  | 0 | 0 | 0 | 1 |  | 0 | 0 | 1 | 0 |
| 47)ООО"АгроСервис-ИТ" | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 48)ООО "Система" | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |
| 49)ООО " веб-студия Умножение" | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |  | 0 | 1 | 0 | 0 |

Кластеризация на меньшее число кластеров делила все компании на две группы по принципу «да», «нет». Кластеризация на большее число кластеров немотивированно относила компании, давшие одинаковые показатели к разным кластерам. Это происходило из-за того, что группы «внутренние процессы компании», «социальные программы и образовательная деятельность» оценивались по шести параметрам, где в основном были бинарные ответы «да», «нет», и это при обработке данных программой не дало разброса по числовой шкале.

В таблице приведены матрицы принадлежности кластерам при кластеризации по группам критериев «внутренние процессы компании», «инновационная деятельность», «социальные программы и образовательная деятельность» при . В таблице обозначена принадлежность предприятий к кластерам по названным критериям, на рисунках приведены интегративные показатели, по которым можно выбрать лидера и аутсайдера, диаграмма рисунков демонстрируют долю предприятий в кластерах.

Центры кластеризации

**Центры кластеризации**

**Центры кластеризации**

**Центры кластеризации**

**Центры кластеризации**

**Центры кластеризации**

**Центры кластеризации**

**Центры кластеризации**

**Центры кластеризации**

**Центры кластеризации**

**Центры кластеризации**

**Центры кластеризации**

Рис. 32. Кластеризация по критерию «Внутренние процессы компании» (нормы центров кластеров, доля ИТ-предприятий, входящих в кластер,

По критерию «внутренние процессы компании» лидирует кластер С3, включающий 8% ИТ-предприятий (Группа Компаний ИТМ, ООО АИСТ, ООО «Телеком.ру»). Для него характерно то, что оцениваемый уровень зрелости процессов компании в 3 раза выше, чем у конкурентов, наличие отдела маркетинга делает компанию более конкурентоспособной.

На втором месте самый многочисленный кластер С2 – 65% предприятий, на третьем месте кластер С1 включающий 23% предприятий. В кластере С1 при анкетировании ИТ-компании не указали наличие стратегии развития, у ИТ-предприятий этого кластера нет сертификации по ISO серии 9000. Аутсайдер – кластер С4 (4% ИТ-предприятий), где кроме этого уровень зрелости процессов ИТ-компании не оценен или ноль.

Центры кластеризации

Рис. 33. Кластеризация по критерию «Инновационная деятельность» (нормы центров кластеров, доля ИТ-предприятий, входящих в кластер,

При заполнении анкеты по критерию «инновационная деятельность» из нескольких вопросов заполненным оказался только один вопрос анкеты «доля инновационных наукоемких продуктов в объеме реализации». По нему производилась кластеризация. 100% инновационных наукоемких продуктов в объеме реализации имеют 4 % ИТ-предприятий (кластер С3: ИП Хомяченко, ООО «Купи батон» ), 25 % инновационных наукоемких продуктов – у 4% ИТ-предприятий (кластер С4: ИнтелСофт, ООО СпецМонтажСервис), 10 % – у 2 % ИТ-предприятий (кластер С3). Остальные ИТ-предприятия (кластер С2), по их оценкам, не производят инновационных наукоемких продуктов.

Центры кластеризации

Рис. 34. Кластеризация по критерию «Социальные программы и образовательная деятельность**»** (нормы центров кластеров, доля ИТ-предприятий, входящих в кластер,

При кластеризации по критерию «социальные программы и образовательная деятельность» первое место занимает кластер С1 (2% предприятий: СимбирСофт), для которого характерно активное участие в организации олимпиад и конкурсов (Волга ИТ 2008-2013, Мастер ИТ 2008-2013); лекций, семинаров и конференций летних, воскресных  школ, лагерей (Летняя школа от УлГТУ, Лесная Летняя школа 2011-13). Объем финансирования социальных программ – 2 % от объема реализации. Ежегодно на базе СимбирСофт проходят практику 8-10 студентов.

На втором месте кластер С3 (4% ИТ-предприятий: Веб-студия Creater, ОООСпецМонтажСервис). Объем финансирования социальных программ –5 % от объема реализации, участие в социальных мероприятиях, проводимых ИТ-сообществом, практикантов нет. Большинство ИТ-предприятий, принадлежащих кластерам С4 (4%), С2 (90%), имеют низкие показатели в социальной сфере.

## § 4.6. Кластеризация по всем группам параметров.

Проанализируем результаты кластеризации предприятий на 6, 8, 10 предприятий по всем группам параметров. В таблице 11 приведены матрицы принадлежности ИТ-предприятий к кластерам.

Кластер С1 – ИП Хомяченко Ю.Н. выделено в отдельный кластер. Данное ИТ-предприятие занимается научными исследованиями, направлением «Умный дом», придерживается стратегии патентного рейдерства, киберсквотинга, открытия филиалов по России по разработке, установке и обслуживанию систем класса «Умный дом». Объем реализации IT-продуктов/услуг за 2012 год составляет до 500 тыс. руб., при этом 100 % доля самофинансирования. Все сотрудники с высшим образованием.

Кластер С2 – ООО «Купи батон» – занимается разработка собственных программных продуктов. Объем реализации IT-продуктов/услуг за 2012 год составил около 10 млн. руб. Это ИТ-предприятие имеет самое большое количество клиентов – около 2,9 млн. человек. По их мнению, конкурентные преимущества и особенности их продукции: «Создав собственную быструю систему синхронизации, позволяющую множеству пользователей одновременно легко управлять общими списками и добавив простой в использовании пользовательский интерфейс мы получили существенную долю рынка и продолжаем успешно наращивать пользовательскую базу завоевывая лидирующие позиции».

Кластер С3 – «Веб-студия Creater». Предприятие «Веб-студия Creater» занимается заказной разработкой программного обеспечения (интернет-сайтов), продвижением интернет-ресурсов (SEO, SMM, и т.п.), разработкой собственных программных продуктов. Количество сотрудников на предприятии в 2012 г. – 16 человек. Объем реализации IT-продуктов/услуг за 2012 год составил около 10 млн. руб.. 150 выполненных проектов в 2012 г. и около сотни клиентов. Средняя цена продукции (услуг) 1000 рублей / 1 час, когда средняя цена продукции (услуг) у конкурентов 1500 рублей / 1 час. Почти все сотрудники с высшим образованием (около 90%).

Кластер С4 – ООО «АктивУчет» – занимается заказной разработкой программного обеспечения (интернет-сайтов), ремонтом и обслуживанием компьютерной техники, сопровождением 1С, обучением IT-технологиям, системным администрированием и настройкой локальных вычислительных систем. В компании, судя по анкетным данным, работает всего один сотрудник с высшим образованием. Объем реализации IT-продуктов/услуг за 2012 г. составил около 500 тыс. руб. Средняя цена продукции (услуг) 800 руб/час, когда у конкурентов средняя цена продукции (услуг) 750-850 руб/час. Есть студенты, проходящие производственную практику на базе предприятия. Сотрудниками ведется кружок информатики в начальной школе.

Кластер С5 – ООО «СпецМонтажСервис» – занимается научными исследованиями, спутниковым Интернетом и телевидением. Всего в компании 6 сотрудников, около 500 клиентов. 100 % – доля самофинансирования. 83 % сотрудников с высшим образованием. 17 % сотрудников – студенты.

При кластеризации на 6 кластеров самым многочисленный кластером оказался кластер C6 в него входят почти 90 % всех предприятий. Данный кластер обладает наибольшим объемом реализации IT-продуктов/услуг за 2012 г., наименьшей долей самофинансирования и наименьшей долей сотрудников с высшим образованием, при этом наибольшим количеством сотрудников.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | С1 | С2 | С3 | С4 | С5 | С6 |  | С1 | С2 | С3 | С4 | С5 | С6 | С7 | С8 |  | С1 | С2 | С3  Таблица 11 | С4 | С5 | С6 | С7 | С8 | С9 | С10 |
| 01)ИнтелСофт | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 02)СимбирСофт | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 03)ИП Хомяченко Ю.Н. | 1 | 0 | 0 | 0 | 0 | 0 |  | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |  | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 04)JackNyfe (Echo) | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 05)Группа Компаний ИТМ | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 06)ООО "Купи батон" | 0 | 1 | 0 | 0 | 0 | 0 |  | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |  | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 07)JackNyfe(Echo) | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 08)ООО АИСТ | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 09)ООО "Телеком.ру" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 10)Буферная бухта | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 11)ITmaster | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 12)Веб-студия Creater | 0 | 0 | 1 | 0 | 0 | 0 |  | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |  | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13)ооо "андер девелопмент" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 14)ITECH.group | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 15)ЗАО "Софткей" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 16)B2B Дизайн-бюро «Зебра» | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 17)ЗАО"РЕГИОН.АТТЕСТ. ЦЕНТР" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 18)ООО"Автоматизация" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 19)ООО"Автоном" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 20)ООО"Волга-Партнер" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 21)ООО "Тайм-софт" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 22)ООО"Техномастер" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 23)ООО "Решение" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 24)ООО «Автоком» | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 25)ОООВнедренческий центр "ПаритетЪ" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 26)ОАО "Ростелеком " | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 27)Группа компаний Симтек | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 28)ООО "Арсенал" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 29)ООО«ХайТэк Девелопмент Групп» | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 30)ООО «ТК «Альфа» | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 31)ООО «Тайм-Софт» | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 32)ООО «Цифровая Бомба» | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 33)ООО "АктивУчет" | 0 | 0 | 0 | 1 | 0 | 0 |  | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 34)ООО «Решение» | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 35)Litota Labs | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 36)ООО "ЗЕВС" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 37)ООО «Тауруна» | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 38)ООО "УмКо" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 39)ООО "Креативная разработка" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 40)ЗАО "Потисс" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 41)1С-Рарус | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 42)ООО"Автоматизация" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 43)ЗАО Ультрамарин | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 44)Планикс | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 45)Artwedis | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 46)ОООСпецМонтажСервис | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 47)ООО"АгроСервис-ИТ" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 48)ООО "Система" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 49)ООО " веб-студия Умножение" | 0 | 0 | 0 | 0 | 0 | 1 |  | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |

Таблица 12

Нормированные значения всех параметров

при кластеризации на 6 кластеров

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | С1 | С2 | С3 | С4 | С5 | С6 |
| 1.1)Область деятельности компании | 0,54545 | 0,03482 | 0,03482 | 0,08027 | 0,68182 | 0,118727 |
| 2.1) Объем реализации IT-продуктов/услуг за год | 0,001 | 0,02 | 0,02 | 0,001 | 0 | 0,049258 |
| 2.2)Объем реализации IT-продуктов за год в других странах | 0 | 0,1 | 0,1 | 0 | 0 | 0,045994 |
| 2.3)Прирост объема реализации за г. | 0,42857 | 1,42857 | 1,14286 | 0,32857 | 0,42857 | 0,43322 |
| 2.6)Доля самофинансирования | 1 | 1 | 1 | 0 | 1 | 0,325938 |
| 3.1)Количество клиентов | 0 | 1 | 0,00003 | 0,00001 | 0,00017 | 0,000178 |
| 3.2)Количество выполненных проектов | 0 | 0,001 | 0,15 | 0 | 0 | 0,032992 |
| 3.7)Уровень удовлетворенности потребителя | 0 | 0,09 | 0,94 | 0 | 0,99 | 0,124186 |
| 3.10)Доля рынка, занимаемая компанией | 0 | 0 | 0 | 0 | 0,23077 | 0,070178 |
| 3.11.1)Средняя цена вашей продукции (услуг) | 0 | 0 | 0,5 | 0,4 | 0 | 0,061777 |
| 3.11.2)Средняя цена продукции у ваших конкурентов рублей / час | 0 | 0 | 0,75 | 0,425 | 0 | 0,072783 |
| 4.1)Средний возраст сотрудников | 0,8 | 0,6 | 0,6 | 1 | 0 | 0,585269 |
| 4.2)Сотрудников с высшим образованием | 1 | 0,9 | 0,9 | 1 | 0,83 | 0,618143 |
| 4.3)Сотрудников, являющихся студентами | 0 | 0,2 | 0 | 0 | 0,34 | 0,115731 |
| 4.4)Количество и сотрудников с научными степенями | 0 | 0,06494 | 0 | 0 | 0 | 0,033404 |
| 4.5)Сотрудников, являющихся выпускниками 2013 года | 0 | 0 | 0 | 0 | 0 | 0,083174 |
| 4.6)Есть ли корпоративное обучение иностранным языкам | 0 | 0 | 0 | 0 | 0 | 0,184241 |
| 4.7)Корпоративное обучение новым технологиям разработки ПО и т.п. | 0 | 0 | 1 | 0 | 0 | 0,363951 |
| 4.8)Количество сотрудников (за 5лет), участвующих в конференциях,… | 0 | 0,35 | 0 | 0,05 | 0 | 0,058949 |
| 4.9)Количество сотрудников | 0,0005 | 0,005 | 0,008 | 0,0005 | 0,003 | 0,035304 |
| 5.1)Стратегия развития компании | 0,01887 | 0 | 0 | 0,01887 | 0,01887 | 0,01513 |
| 5.2)Возраст компании | 0,00867 | 0,00372 | 0,01363 | 0,01363 | 0,01363 | 0,011941 |
| 5.5)Сертификация по ISO серии 9000 | 0 | 0 | 0 | 0 | 0 | 0,023188 |
| 5.6)Оцениваемый уровень зрелости процессов компании | 0 | 0 | 0 | 0 | 0 | 0,019327 |
| 5.7)Наличие отдела маркетинга | 0 | 0,03704 | 0,03704 | 0 | 0,03704 | 0,017747 |
| 5.8)Наличие отдела продаж | 0 | 0,02564 | 0,02564 | 0 | 0,02564 | 0,016801 |
| 6.1)Доля инновационных наукоемких продуктов в объеме реализации | 1 | 1 | 0 | 0 | 0,2 | 0,008421 |
| 7.1)Объем финансирования соц. программ в % от объема реализации | 0 | 0,2 | 1 | 0 | 1 | 0,013517 |
| 7.2)Организация, участие компании в проведении олимпиад и конкурсов | 0 | 1 | 1 | 0 | 0 | 0,17862 |
| 7.3)Организация лекций, семинаров и конференций | 0 | 1 | 1 | 0 | 1 | 0,317178 |
| 7.4)Организация летних, воскресных  школ, лагерей | 0 | 0 | 1 | 0 | 0 | 0,155909 |
| 7.5)Кол-во студентов, проходящих практику на  базе предприятия | 0 | 0,00053 | 0 | 0,025 | 0 | 0,055308 |
| 7.11)Количество дисциплин преподаваемых сотрудниками | 0 | 0 | 0 | 0 | 0 | 0,056408 |

В столбцах таблицы 12 приведены координаты векторов – центров кластеров при кластеризации на 6 кластеров по всем группам параметров. Нормировка векторов произведена так, что каждая координата обозначает долю от предприятия, имеющего максимальное значение соответствующего параметра. Для центра кластера берется средняя по кластеру величина показателя.

Центрыкластеризации

Норма

Норма

Центры кластеризации

Норма

Рис. 35. Кластеризация по всем критериям (нормы центров кластеров)

Если судить по графикам центров кластеризации, то на первом месте всегда оказывается кластер, состоящий из предприятия «Веб-студия Creater», которое гармонично представлено по всем группам параметров. На втором месте оказывается кластер, состоящий из предприятия ООО «Купи батон», выделившийся числом клиентов. На третьем месте оказывается кластер, состоящий из предприятия ООО «СпецМонтажСервис», которое занимается научные исследованиями, осуществляет деятельность на рынке услуг спутниковых операторов ТВ, осуществляет монтаж СКС, охранного видеонаблюдения, монтаж ОПС.

## § 4.7. Кластеризация ИТ-предприятий Ульяновской области FCM-методом. Проверка надежности и достоверности результатов кластеризации.

Была проведена кластеризация ИТ-предприятий Ульяновской области с использование алгоритма нечетких -средних (FCM).

Алгоритм FCM по своему характеру относится к приближенным алгоритмам поиска экстремума для целевой функции при наличии ограничений. Поэтому в результате выполнения данного алгоритма определяется локально-оптимальное нечеткое разбиение, которое описывается совокупностью функций принадлежности, а также центрами или типичными представители каждого из нечетких кластеров.

В качестве метода нечеткой кластеризации выберем FCM, число кластеров равное 6, количество итераций – 20, параметр, отвечающий за «нечеткость» кластеризации . Чем больше , тем более нечеткой является кластеризация. Для сравнения на рисунке 25 приведен фрагмент таблицы при . В таблице 13 приведены результаты кластеризации FCM-методом.

Чем ближе значение функции принадлежности в таблице 13 к 1, тем с большим основанием ИТ-предприятие может быть отнесено к данному кластеру. В таблице 13 выделим цветом клетки с максимальными показателями в строке.

Сравним таблицу 11 и таблицу 13 с результатами кластеризации разными методами по всем параметрам. Заключаем, что в основной группе (в таблице 11 кластер С6, в таблице 13 – кластер С5) в 77% случаев результаты кластеризации совпадают.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | C1 | C2 | C3 | C4  Таблица 13 | C5 | C6 |
| 01)ИнтелСофт | 0,001 | 0,933 | 0,010 | 0,000 | 0,055 | 0,000 |
| 02)СимбирСофт | 0,034 | 0,295 | 0,205 | 0,025 | 0,389 | 0,052 |
| 03)ИП Хомяченко Ю.Н. | 0,017 | 0,365 | 0,388 | 0,024 | 0,195 | 0,011 |
| 04)JackNyfe (Echo) | 0,000 | 0,001 | 0,000 | 0,000 | 0,998 | 0,000 |
| 05)Группа Компаний ИТМ | 0,000 | 0,024 | 0,010 | 0,000 | 0,965 | 0,000 |
| 06)ООО "Купи батон" | 0,125 | 0,233 | 0,211 | 0,112 | 0,218 | 0,102 |
| 07)JackNyfe(Echo) | 0,004 | 0,133 | 0,068 | 0,007 | 0,787 | 0,001 |
| 08)ООО АИСТ | 0,000 | 0,000 | 0,000 | 0,000 | 1,000 | 0,000 |
| 09)ООО "Телеком.ру" | 0,000 | 0,003 | 0,001 | 0,000 | 0,996 | 0,000 |
| 10)Буферная бухта | 0,000 | 0,003 | 0,001 | 0,000 | 0,996 | 0,000 |
| 11)ITmaster | 0,000 | 0,000 | 0,000 | 0,000 | 0,999 | 0,000 |
| 12)Веб-студия Creater | 0,023 | 0,047 | 0,042 | 0,021 | 0,047 | 0,819 |
| 13)ооо "андер девелопмент" | 0,004 | 0,081 | 0,016 | 0,002 | 0,897 | 0,001 |
| 14)ITECH.group | 0,000 | 0,000 | 0,000 | 1,000 | 0,000 | 0,000 |
| 15)ЗАО "Софткей" | 0,000 | 0,024 | 0,009 | 0,000 | 0,965 | 0,000 |
| 16)B2B Дизайн-бюро «Зебра» | 1,000 | 0,000 | 0,000 | 0,000 | 0,000 | 0,000 |
| 17)ЗАО"РЕГИОН.АТТЕСТ. ЦЕНТР" | 0,005 | 0,859 | 0,041 | 0,004 | 0,089 | 0,002 |
| 18)ООО"Автоматизация" | 0,000 | 0,184 | 0,011 | 0,000 | 0,804 | 0,000 |
| 19)ООО"Автоном" | 0,000 | 0,000 | 0,000 | 0,000 | 1,000 | 0,000 |
| 20)ООО"Волга-Партнер" | 0,000 | 0,000 | 0,000 | 0,000 | 1,000 | 0,000 |
| 21)ООО "Тайм-софт" | 0,000 | 0,000 | 0,000 | 0,000 | 1,000 | 0,000 |
| 22)ООО"Техномастер" | 0,000 | 0,002 | 0,000 | 0,000 | 0,998 | 0,000 |
| 23)ООО "Решение" | 0,000 | 0,959 | 0,005 | 0,000 | 0,036 | 0,000 |
| 24)ООО «Автоком» | 0,000 | 0,000 | 0,000 | 0,000 | 0,999 | 0,000 |
| 25)ОООВнедренческий центр "ПаритетЪ" | 0,000 | 0,959 | 0,005 | 0,000 | 0,036 | 0,000 |
| 26)ОАО "Ростелеком " | 0,004 | 0,126 | 0,016 | 0,001 | 0,852 | 0,001 |
| 27)Группа компаний Симтек | 0,000 | 0,001 | 0,000 | 0,000 | 0,999 | 0,000 |
| 28)ООО "Арсенал" | 0,000 | 0,002 | 0,000 | 0,000 | 0,998 | 0,000 |
| 29)ООО «ХайТэк Девелопмент Групп» | 0,000 | 0,000 | 0,000 | 0,000 | 0,999 | 0,000 |
| 30)ООО «ТК «Альфа» | 0,000 | 0,000 | 0,000 | 1,000 | 0,000 | 0,000 |
| 31)ООО «Тайм-Софт» | 0,000 | 0,001 | 0,000 | 0,000 | 0,999 | 0,000 |
| 32)ООО «Цифровая Бомба» | 0,007 | 0,220 | 0,107 | 0,005 | 0,656 | 0,004 |
| 33)ООО "АктивУчет" | 0,003 | 0,052 | 0,877 | 0,006 | 0,059 | 0,002 |
| 34)ООО «Решение» | 0,000 | 0,001 | 0,997 | 0,000 | 0,002 | 0,000 |
| 35)Litota Labs | 0,000 | 0,001 | 0,000 | 0,000 | 0,999 | 0,000 |
| 36)ООО "ЗЕВС" | 0,000 | 0,001 | 0,000 | 0,000 | 0,999 | 0,000 |
| 37)ООО «Тауруна» | 0,000 | 0,000 | 0,000 | 0,000 | 0,999 | 0,000 |
| 38)ООО "УмКо" | 0,000 | 0,000 | 0,000 | 0,000 | 1,000 | 0,000 |
| 39)ООО "Креативная разработка" | 0,000 | 0,003 | 0,990 | 0,000 | 0,006 | 0,000 |
| 40)ЗАО "Потисс" | 0,000 | 0,004 | 0,991 | 0,000 | 0,005 | 0,000 |
| 41)1С-Рарус | 0,000 | 0,002 | 0,000 | 0,000 | 0,998 | 0,000 |
| 42)ООО"Автоматизация" | 0,000 | 0,000 | 0,000 | 0,000 | 1,000 | 0,000 |
| 43)ЗАО Ультрамарин | 0,000 | 0,004 | 0,001 | 0,000 | 0,995 | 0,000 |
| 44)Планикс | 1,000 | 0,000 | 0,000 | 0,000 | 0,000 | 0,000 |
| 45)Artwedis | 0,000 | 0,003 | 0,990 | 0,000 | 0,006 | 0,000 |
| 46)ОООСпецМонтажСервис | 0,029 | 0,016 | 0,010 | 0,009 | 0,015 | 0,922 |
| 47)ООО"АгроСервис-ИТ" | 0,000 | 0,000 | 0,000 | 0,000 | 0,999 | 0,000 |
| 48)ООО "Система" | 0,000 | 0,000 | 0,000 | 0,000 | 1,000 | 0,000 |
| 49)ООО " веб-студия Умножение" | 0,000 | 0,001 | 0,000 | 0,000 | 0,999 | 0,000 |

ИТ-предприятий, попавших в другие кластеры, стало больше. Здесь результаты кластеризации FCM-методом более сопоставимы с более «дифференцированной» кластеризацией иерархическим методом при (см. таблицу 11). ИТ-предприятия, выделившиеся в другие кластеры, в основном, – те же.

В С2 вошли ИнтелСофт, ИП Хомяченко, ООО «Купи батон», ЗАО «РЕГИОН.АТТЕСТ. ЦЕНТР», ООО «Решение», ООО Внедренческий центр "ПаритетЪ", большинство из которых выделялось в отдельные кластеры при иерархическом методе. В кластер С3 вошли ИП Хомяченко, ООО «Купи батон», ООО «АктивУчет», ООО «Решение», ООО «Креативная разработка», ЗАО «Потисс», Artwedis. В кластер С4 – ITECH.group, ООО «ТК «Альфа». В С6 – Веб-студия Creater, ОООСпецМонтажСервис. Необходимо отметить, что большинство из этих предприятий – лидеры в кластеризации иерархическим методом по отдельным группам параметров (см. таблицу 14), т.е. выделение их в отдельные кластеры FCM-методом неслучайно.

Точного совпадения кластеров, кроме самого многочисленного, нет, но все лидеры и аутсайдеры иерархического метода по различным группам параметров «засветились» FCM-методом.

Отсюда видно, что метод FCM более «тонкий» метод кластеризации, нежели центроидный метод. Его достоинства – простота и быстрота использования; нечеткость при определении объекта в кластер, что позволяет определять объекты, которые находятся на границе, в кластеры.

В таблице 14 приведен сравнительный анализ кластеризации при центроидным методом для каждой группы параметров отдельно и для всех параметров.

Таблица 13

Кластеризация по отдельным показателям, по всем параметрам

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| кластеры | Область деятельности | Финансово-экономические показатели | Кадровые показатели | Рынок сбыта | Внутренние процессы компании | Инновационная деятельность | Социальные программы и образовательная деятельность | Кластеризация по всем параметрам |
| Первый | ЗАО"РЕГИОНАЛЬНЫЙ АТТЕСТАЦИОННЫЙ ЦЕНТР" **ОООСпецМонтажСервис** | JackNyfe(Echo) ITECH.group,  ООО «ТК «Альфа» | ООО "АктивУчет" | **ООО "Купи батон"** | Группа Компаний ИТМ, ООО АИСТ, ООО "Телеком.ру" ,ITmaster | ИП Хомяченко, **ООО "Купи батон"** | СимбирСофт , **Веб-студия Creater,** **ОООСпецМонтажСервис** | **Веб-студия Creater,**  **ООО "Купи батон"** |
| Второй | ИнтелСофт  **ИП** **Хомяченко** | остальные | **ИП Хомяченко**, ИнтелСофт, ITECH.group | Остальные | Остальные | **ОООСпецМонтажСервис,** ИнтелСофт | ООО «Цифровая Бомба», **ООО "Купи батон"** | **ОООСпецМонтажСервис, Хомяченко Ю.Н.** |
| Третий | Остальные | Группа Компаний ИТМ | Остальные | Планикс | ИнтелСофт ,СимбирСофт, ИП Хомяченко, JackNyfe (Echo) , "Купи батон", , Буферная бухта , Веб-студия, Creater, «ТК «Альфа», ООО «Тайм-Софт», ООО «Цифровая Бомба», "ЗЕВС", ЗАО Ультрамарин | Остальные | Остальные | Остальные |

Первый кластер – самый отдаленный от начала координат, т.е. он характеризуется более высокими показателями, ИТ-предприятия попавшие в него можно назвать лидерами. У третьего кластера самые низкие показатели. В группе «Область деятельности» наибольшее расстояние центра кластера от начала координат означает редкость соответствующей сферы деятельности. В группе «Финансово-экономические показатели» данные и результаты кластеризации искажены тем, что многие предприятия не указали объемов реализации, и программа автоматических считала их равными нулю.

Обращает на себя внимание, что центроидный метод кластеризации выбрасывает по каждой группе показателей в отдельные кластеры те ИТ-предприятия, которые значительно отличаются соответствующими показателями от общей массы. Т.е. предъявляется «общая масса» и «белые вороны» – лидеры или аутсайдеры.

При сравнении результатов кластеризации по отдельным группам показателей и по всем показателям видно, что лидеры отдельных групп становятся лидерами кластеризации по всем параметрам, т.е. их «лица необщее выражение» сказывает и на общих результатах кластеризации.

**Для оценки надежности и достоверности работы программы** использовались следующие процедуры, обеспечивающие адекватную проверку качества кластерного анализа.

– Кластерный анализ выполнялся на основании одних и тех же данных, но с использованием различных способов измерения расстояния (метрики и ).

– Кластеризация на разное количество кластеров даёт сопоставимые результаты.

– Кластеризация по разным группам и по всем параметрам даёт сопоставимые результаты, лидеры отдельных групп становятся лидерами при общей кластеризации.

– Результаты кластерного анализа по сокращенному и по полному набору переменных сравнимы.

– Результаты, полученные центроидным методом и FCM-методом, сопоставимы.

## Выводы.

Анализ результатов кластеризации показал, что типичная область деятельности ИТ-предприятий Ульяновской области – заказная разработка программного обеспечения: интернет-сайты, веб-приложения, разработка и сопровождение программных продуктов, интернет-провайдеры, ремонт и обслуживание техники, продажа, системная интеграция и администрирование. В Ульяновской области мало ИТ-предприятий занимаются обучением, научными исследованиями в сфере ИТ, защитой информации. Доля наукоёмких продуктов невелика. Лишь 4% предприятий имеют 100%, 4% предприятий – 25%, 2% – 10% наукоёмких продуктов.

В основном ИТ-предприятия малочисленны от 1 до 10 сотрудников. Примерно четверть предприятий имеет боле 100 сотрудников. Средний возраст сотрудников 25-30 лет. Около 80% из них с высшим образованием. Малая доля предприятий занимается обучением своих сотрудников.

В отрасли высокая доля самофинансирования, что связано со спецификой ИТ-отрасли. Средний объём реализации продуктов по самому многочисленному кластеру невысок, 200 – 500 тыс. рублей в год. Это связанно с особенностями заполнения анкеты, где многие предприятия из соображений коммерческой тайны не указали объёмов реализации. Объёмы реализации у предприятий - лидеров – 100 – 500 млн. рублей в год.

Единичные предприятия занимаются социальными программами. Всего у 4 предприятий доля финансирования социальных программ 2 – 5 % от реализации.

# Заключение.

При выполнении работы изучены основы кластерного анализа, элементы теории нечетких множеств, экономические модели кластеризации, произведен анализ критериев и параметров, характеризующих ИТ-предприятия. Изучены возможности и особенности программных продуктов для кластерного анализа..

Осуществлена разработка программного продукта «Кластеризация IT-предприятий», предназначенного для обработки экономических данных и их анализа, группировки, распознавания и представления в удобном для пользователя виде. Наряду с обработкой числовых данных возможна обработка нечетких данных.

Программный продукт был представлен на конкурсе инновационных проектов по программе Фонда содействия развитию малых форм предприятий в научно-технической сфере «Участник молодежного научно-инновационного конкурса 2013» («УМНИК»).

Работа представлена на XXIX международной научно-практической конференции «Инновации в науке» (Новосибирск, 2014).

Опубликована статья Афанасьевой Т.В., Сибирева И.В. «Программа «Сегментации и кластеризация рынка IT» [].

Осуществлена экспериментальная часть работы: кластеризация ИТ-предприятий Ульяновской области по анкетным данным 2012г., произведен анализ кластеризации, текущего уровня развития ИТ-сектора в регионе.

Дальнейшее развитие проекта видится в разработке web-версии программы, нечетком вводе и выводе данных.

# Литература

1. Айвазян, С.А. Прикладная статистика: Классификация и снижение размерности / С.А. Айвазян, В.М. Бухштабер, И.С. Енюков, Л.Д. Мешалкин. – М.: Финансы и статистика, 1989. – 607с.
2. Асаи, К. Прикладные нечеткие системы / К. Асаи, Д. Ватада, С. Иваи и др. / Под ред. Т. Тэрано, К. Асаи, М. Сугено. – М.: Мир, 1993. – 368 с.
3. Афанасьева, Т. В. Нечёткое моделирование временных рядов и анализ нечётких тенденций / Т. В. Афанасьева, Н. Г. Ярушкина. – Ульяновск : УлГТУ, 2009. – 299 с.
4. Афанасьева, Т.В. Программа «Сегментация и кластеризация рынка IT» / Т.В. Афанасьева, В.В. Сибирев // Инновации в науке/ сб.ст. по материалам XXIX междунар. науч.-практ.конф. № 1 (26). –Новосибирск: Изд. «СибАК», 2014. С. 46-53.
5. **Беллман, Р.** Принятие решений в расплывчатых условиях.- В кн.: Вопросы анализа и процедуры принятия решений // **Р. Беллман, Л. Заде**. – М.: Мир, 1976. - С. 172-215.
6. Болч, Б. Многомерные статистические методы для экономики / Б. Болч, К. Дж. Хуань. / Пер. с англ. – М.: Статистика, 1979. – 317с.
7. Бююль, А. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей / А. Бююль. –СПб.: ДиаСофтЮП, 2002, - 608 с.
8. Дубров, А.М. Многомерные статистические методы / А.М. Дубров, В.С. Мхитарян, Л.И. Трошин. –М.: Финансы и статистка, 2000. - 352 с.
9. Дюран, Б. Кластерный анализ / Б.Дюран, П.Оделл. – М.: Статистика, – 1977. – 128 с.
10. Жамбю, М. Иерархический кластер-анализ и соответствия. Пер. с фр. / М. Жамбю. – М.: Финансы и статистика, 1988. – 342 с.
11. Ким, Дж. О. Факторный, дискриминантный и кластерный анализ: Пер с англ./Дж. О. Ким, Ч.У. Мьюллер, У.Р. Клекка и др.; Под ред. И.С. Енюкова. – М.: Финансы и статистика, 1989. – 215с.
12. Классификация и кластер. /Под ред. Дж. Вэн Райзина. – М.: Мир, 1980, –390 с.
13. **Кофман, А.** Введение в теорию нечетких множеств / **А. Кофман**. – М.: Радио и связь, 1982.- 432 с.
14. Леоненков, А.В. Нечеткое моделирование в среде MATLAB и fuzzyTECH. / А.В. Леоненков. – Санкт-Петербург: БХВ-Петербург, 2005. – 736 с.
15. Мандель, И.Д. Кластерный анализ / И.Д. Мандель. – М.: Финансы и статистика. 1988. – 176с.
16. **Недосекин, А.О.** Нечетко-множественный анализ риска фондовых инвестиций / **А.О. Недосекин**. –СПб: Изд-во Сезам, 2002. - 181 с.
17. Портер, М. Международная конкуренция. Конкурентные преимущества стран / М. Портер. – М.: Междунар. отношения, 1993. – 896 с.
18. Сошникова, Л.А. Многомерный статистический анализ в экономике / Л.А. Сошникова, В.Н. Тамашевич, Г. Уебе, М. Шефер. – М.: ЮНИТИ-ДАНА, 1999. – 598 с.
19. Тарасенко, Ф.П. Прикладной системный анализ. (Наука и искусство решения проблем): Учебник / Ф.П. Тарасенко. – Томск,: Изд-во Том.ун-та, 2004, – 186 с.
20. Типология и классификация в социологических исследованиях. Отв. ред. В.Г. Андреенков, Ю.Н. Толстова. –М.: Наука, 1982. – 296с.
21. Ярушкина, Н. Г. Интеллектуальный анализ временных рядов / Н.Г. Ярушкина, Т.В. Афанасьева. – Ульяновск: УлГТУ, 2010.– 299с.
22. Ярушкина, Н. Г. Исследование ИТ-кластера Ульяновской области / Н. Г. Ярушкина, Т. В. Афанасьева, О. В. Шиняева и др., отв. ред. Т. В. Афанасьева. – Ульяновск : УлГТУ, 2013. – 137 с.
23. Ярушкина, Н.Г. Основы теории нечетких и гибридных систем. Учебное пособие / Н.Г. Ярушкина. – СПБ.: Финансы и статистика, 2004. – 320c.
24. Sokal R. And P.Sneat (1963) Principles of Numerical Taxonomy. San Francisco: W.H.Freeman
25. Tryon R.C. Cluster Analysis. New York: McGraw-Hill. - 1939.
26. «ClusterDelta  – кластерный график как инструмент анализа рынка» http://tempofox.com/klasternyj-grafik-kak-instrument-analiza-rynka/
27. «Сегментация рынка – ​​​​​​​ позиционирование товара» <http://www.segmentation-market.ru/index.html>) позволяет производить:
28. «StatisticalPackagefortheSocialSciences» (SPSSStatistics)
29. http://www.datuapstrade.lv/rus/spss/section\_20/2/ –
30. [www.gks.ru/bgd/regl/b12\_14p/Main.htm](http://www.gks.ru/bgd/regl/b12_14p/Main.htm)
31. <http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/science_and_innovations/it_technology/>
32. http://www.hse.ru/primarydata/rir2012ФГУП ВНИИ ПВТИ) http://pvti.ru/)
33. http://www.apkit.ru/files/Strategy\_APKIT\_2012\_vr.pdf
34. <http://www.russoft.ru/reports>
35. <http://eregion.ru/>
36. http://www.tadviser.ru/
37. <http://www.raexpert.ru/>
38. <http://www.cnews.ru/reviews/new/rynok_it_itogi_2012/>
39. <http://www.ratingruneta.ru/>
40. <http://www.cmsmagazine.ru/>
41. http://raec.ru/analytics/
42. http://2012.tagline.ru/
43. <http://www.russoft.ru/reports>
44. Аналитический доклад 2012 г. ассоциации АП КИТ http://www.apkit.ru/files/Strategy\_APKIT\_2012\_vr.pdf
45. «О мерах по развитию отрасли ИТ в Российской Федерации» Мониторингу региональной информатизации и прогнозы (73 субъекта РФ) <http://pvti.ru/monitoring-materials.htm>
46. Сборник публикаций «Индикаторы информационного общества: 2013» http://www.hse.ru/primarydata/iio2013
47. http://www.ratingruneta.ru/research/web2010
48. <http://www.raexpert.ru/>
49. «Показатели развития информационного общества» http://www.gks.ru/wps/wcm/connect/rosstat\_main/rosstat/ru/statistics/science\_and\_innovations/it\_technology/.

## Приложение

## Приложение 1

Таблица

Сравнительный анализ программных пакетов для кластеризации.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Методы | Программная реализация | Функции | Ограничения |
| ClusterDelta | – | Нет доступа | Построение кластерного графика для вычисления точечного объема заданных сделок в ограниченном диапазоне времени и финансов. | Кластерный анализ используется только для построения «ценового бара», вычисления точечного объема заданных сделок в ограниченном диапазоне времени и финансов.  Узкопрофессиональная направленность программы. |
| Сегментация рынка позиционирование товара | «Метод K-means» на основе меры сходства | Нет доступа | 1.​ Выделение сегментов рынка методами экспертного сегментирования рынка.  2.​ Выделение целевых сегментов рынка методами кластерного анализа (К- means алгоритмами).  3.​ При оценке сходства потребителей учитывается важность отдельных переменных.  4.​ Анализ перекрытия выделяемых сегментов рынка. Сегментация рынка  5.​ Сегментация рынка - выделение и интерпретация семантики выделенных сегментов.  6.​ Исследование выделенных сегментов с помощью функций полезности, главных факторов, профилей сегментов при сегментации рынка.  7.​ Позиционирование товаров и бренда на выделенных сегментах. Построение карты восприятия товара (карта позиционирования).  8.​ Исследование стратегического потенциала целевых сегментов рынка (портфолио-анализ).  9.​ Представление результатов анализа доходных сегментов рынка в виде таблиц и графиков. | Исследуется рынок товаров |
| SPSS Statistics | 1.​ Межгрупповые связи.  2.​ Связь внутри групп.  3.​ Близлежащий сосед.  4.​ Дальний сосед.  5.​ Центроидная кластеризация.  6.​ Медианная кластеризация.  7. Метод Варда. | Нет доступа | Коммерческих статистический пакет для социальных наук. | Возможно применение профессиональными математиками.  Нет конкретизации терминологии по предметным областям.  Не предусмотрена работа с нечеткими данными и нечеткой кластеризацией. |
| MatLab | Методы иерархического кластерного анализа:   * метод ближнего соседа; * метод дальнего соседа; * метод средней связи; * центроидный метод; * метод медианной связи.   Неиерархический метод кластерного анализа (K-means).  Нечеткая неиерархическая кластеризация (fcm-метод) | Доступ есть | MatLab - система компьютерной м3атематики предназначена для выполнения инженерных и научных вычислений.  Язык MatLab является высокоуровневым интерпретируемым языком программирования. | Реализуются с помощью программирования с использованием встроенных функций.  Применение не программистами затруднительно |