§ 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.

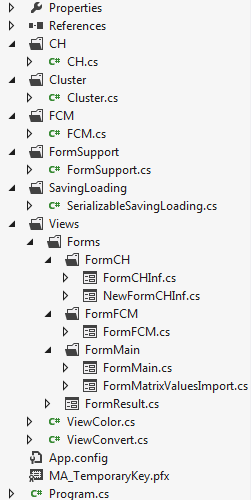


Рис. 26. Структура проекта программы

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.