

NEWS

OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN

PHYSICO-MATHEMATICAL SERIES

ISSN 1991-346X

Volume 2, Number 306 (2016), 123–127

**INTEGRATION OF R SOFTWARE ENVIRONMENT
IN C# SOFTWARE ENVIRONMENT****N. Litvinenko**Institute of mathematics and mathematical modeling MES RK, Almaty, Kazakhstan
e-mail: n.litvinenko@inbox.ru

Key words: R software environment, mathematical model, software environment, environment integration, rscproxy, rcom, software suite (SS).

Abstract. One of the main trend in computer modeling of applied problems in different areas of science and industry is the integration of software environments. This problem is not new, but its importance is increasing all the time. The most common software suite, such as MATLAB, MAPLE, R, SPSS include a huge number of good, well-established, fast libraries and other components, covering a variety of areas in solving of mathematical problems, statistical processing of data, processing large amounts of data, etc. in different software environments. In the process of software modules developing it would be appropriate to use the existing components of these packages. However, the using of these packages is rather difficult due to a number of reasons. In this article we will describe the basic principles of integration of the some modules of the R software environment in the C# software environment. The R SS is a free software product designed for statistical data processing. Today R SS contains more than 7700 packages, encompassing a variety of areas in statistics.

This method will be used in the development of the grant project «Automated techniques for social-psychological diagnostics of military teams».

УДК 004.43; 519.254

ИНТЕГРАЦИЯ ПРОГРАММНОЙ СРЕДЫ R В СРЕДУ C#**Н. Литвиненко**

Институт математики и математического моделирования КН МОН РК, Алматы, Казахстан

Ключевые слова: среда R, математическая модель, среда разработки, интеграция среды, rscproxy, rcom. Пакет прикладных программ (ППП).

Абстракт. Одним из основных научных и прикладных направлений, возникающих при компьютерном моделировании прикладных задач в различных областях науки и производства является интеграция компьютерных вычислительных сред. Данная проблема не нова, однако ее актуальность все время возрастает. Наиболее распространенные пакеты прикладных программ, такие как, например, MATLAB, MAPLE, R, SPSS имеют огромное количество прекрасных, хорошо отлаженных, быстро работающих библиотек и других компонент, охватывающих самые разнообразные направления при решении математических задач, статистической обработке данных, обработке больших объемов данных и пр. в различных вычислительных средах. При разработке программных модулей, являющихся частью разрабатываемых проектов, было бы целесообразно использовать уже готовые компоненты из данных пакетов. Однако задача использования сторонних пакетов достаточно сложна в силу ряда причин. В данной статье описываются основные принципы интеграции отдельных модулей пакета R в среду C#. ППП R является свободно распространяемым программным продуктом, ориентированным на статистическую обработку данных. В настоящее время ППП R содержит более 7700 пакетов, охватывающих самые разнообразные направления в статистике.

Данный способ будет использован при разработке грантового проекта «Автоматизированная методика социально-психологической диагностики воинского коллектива».

Описание проблемы. При разработке грантового проекта «Автоматизированная методика социально-психологической диагностики воинского коллектива» возникает много задач по статистической обработке данных анкетирования. Проект оформляется в виде программного продукта, ориентированного на использование рядовыми социологами. Программный продукт разрабатывается в среде C#. Предполагается в качестве дополнительных необязательных опций привлечение отдельных модулей из свободного ППП R. В данной статье описываются основные принципы интеграции ППП R в среду C#. Интеграция рассматривается на простом учебном примере, никакого отношения к проекту не имеющем. На рисунке 1 приведен внешний вид учебного примера. Далее приведена часть программы, показывающая принцип работы ППП R в среде C#.

При исследовании прикладной проблемы, вначале необходимо построить математическую модель исследуемого процесса. При этом обычно решаются следующие задачи:

1. Выделить суть изучаемого процесса.
2. Построить математическую модель, адекватно отражающую суть изучаемого процесса.
3. Разбить изучаемую проблему на отдельные максимально независимые друг от друга задачи.
4. Определить математические методы решения данных задач.
5. Определить наиболее соответствующую программную среду для решения данных задач.

В данной статье рассматриваются некоторые вопросы 5 задачи - метод интеграции среды R в среду C#. Вопросы инсталляции программной среды R в среды Windows, Linux, MacOS хорошо описаны в технической документации [1-2]. В данном примере мы рассматриваем инсталляцию R только в среду Windows. Экспорт и импорт данных из среды R в среды Excel, Word, Matlab хорошо описаны в технической документации [3]. Данный документ будет настольной книгой при использовании среды R в разрабатываемом проекте. Стоит заметить, что экспорт и импорт данных задача гораздо более простая чем интеграция всей среды R в среду C#. Идеология разработки среды R ориентирована на среду C++. Вопросы интеграции в среду C# рассмотрены крайне слабо. Однако среда C# в .NET является в настоящее время наиболее востребованной средой в разработке программного обеспечения. В технической документации [4] описываются принципы интеграции среды R в среды C++, Excel, Word, Matlab. Мы использовали данную документацию для выработки общей идеологии интеграции среды R в среду .NET (C#).

Данная статья пишется по работе над грантовой темой «Автоматизированная методика социально-психологической диагностики воинского коллектива». Одной из важных задач данной темы является статистическая обработка массива анкет. Основные статистические методы обработки данных описаны в [5]. Данный документ использовался при построении тестового примера. Многие методы обработки анкет хорошо разработаны в вычислительной среде R. Пакет программ для грантовой темы разрабатывается в среде C#. Весьма интересно использовать готовые общедоступные разработки среды R в разрабатываемом проекте, не выходя при этом из среды C#. В книге [6] приведено множество примеров для первоначального изучения среды R. Некоторые примеры из этой книги также используются в тестовом примере. Настольной книгой при написании тестовой программы на языке C# для нас являлась книга [7], но при желании можно использовать и другие.

Среда разработки. Данная работа выполнялась в среде:

- Материнская плата – Gigabyte Technology Co. Ltd., Z77MX-D3H
- Процессор – Intel(R) Core(TM) i7-3770 CPU@ 3.40GHz 3.90GHz
- Память - 16 Гб
- Внешняя память – 2 Тб
- Видеокарта – NVIDIA GeForce GTX 660
- Операционная система – Windows 7 Ultimate, 32 bit
- Среда разработки – Microsoft Visual Studio Professional 2013, версия 12.0.21005.1 REL

Интеграция среды R в среду C#. Почему выгодно использовать среду R?

- R является Свободным программным обеспечением

- R — пакет, в основном ориентированный на разнообразные задачи статистического анализа. Однако данный пакет может решать и многие другие задачи.

- R используется в социальных науках, статистике, экономике, страховом деле, социологии, финансах и т.д.

- R доступен для основных операционных систем Linux, Mac OS, Windows

- Количество общедоступных прикладных пакетов на 1 января 2015 превышает 7700 только на официальном сайте CRAN.

- R — объектно-ориентированный язык.

- R — матричный язык.

- Каждый пакет R строится стандартным образом, имеет хорошую поддержку, многие пакеты содержат учебные примеры.

Интеграция пакета R в среду C#:

1. Вначале устанавливаем последнюю версию R-3.2.3-win.exe и проводим стандартную процедуру установки пакета R.

2. Запускаем R.

3. В среде R устанавливаем пакет rscproху

```
>install.packages("rscproху")
```

4. В среде R устанавливаем пакет rcom

```
>install.packages("rcom")
```

5. Скачиваем пакет R_scilab_DCOM3.0-1B5 по ссылке http://www.freestatistics.org/cran/contrib/extra/dcom/R_Scilab_DCOM3.0-1B5.exe и устанавливаем его.

6. Запускаем Visual Studio 2013. Создаем тестовый пример в приложении Windows Forms в среде C# с именем Bilbak_R_C_3

7. Стандартная форма тестового примера содержит меню, которое позволяет запустить некоторые примеры из пакетов R. На форме находятся 4 текстовых окна. В первом окне показано как выглядят команды текущего примера в среде R. Во втором окне показаны команды для этого же примера в среде C#. В третьем окне выводятся результаты. В четвертом окне показаны краткие пояснения по текущему примеру. В основном это пояснения по синтаксису команд среды R.

8. Для корректной работы тестовых программ обеспечиваем ссылки на следующие библиотеки: StatConnectorCommonLib, STATCONNECTORSRVLib, StatConnTools. Данные библиотеки образуются при установке пакета R_scilab_DCOM3.0-1B5.

9. Для подключения ссылок необходимо сделать следующее: в обозревателе решений, правой клавишей мыши по References выбираем пункт меню «добавить ссылку», в открывшемся окне выбираем COM-библиотеки типов, из открывшегося списка выбираем интересующие нас библиотеки из пункта 8.

10. Для корректной работы необходимо также в начале текста программы указать соответственно:

```
using StatConnectorCommonLib;
```

```
using STATCONNECTORSRVLib;
```

```
using StatConnTools;
```

11. На рисунке 1 приведен внешний вид тестовой программы.

12. После рисунка 1 приведена часть программного кода тестовой программы.

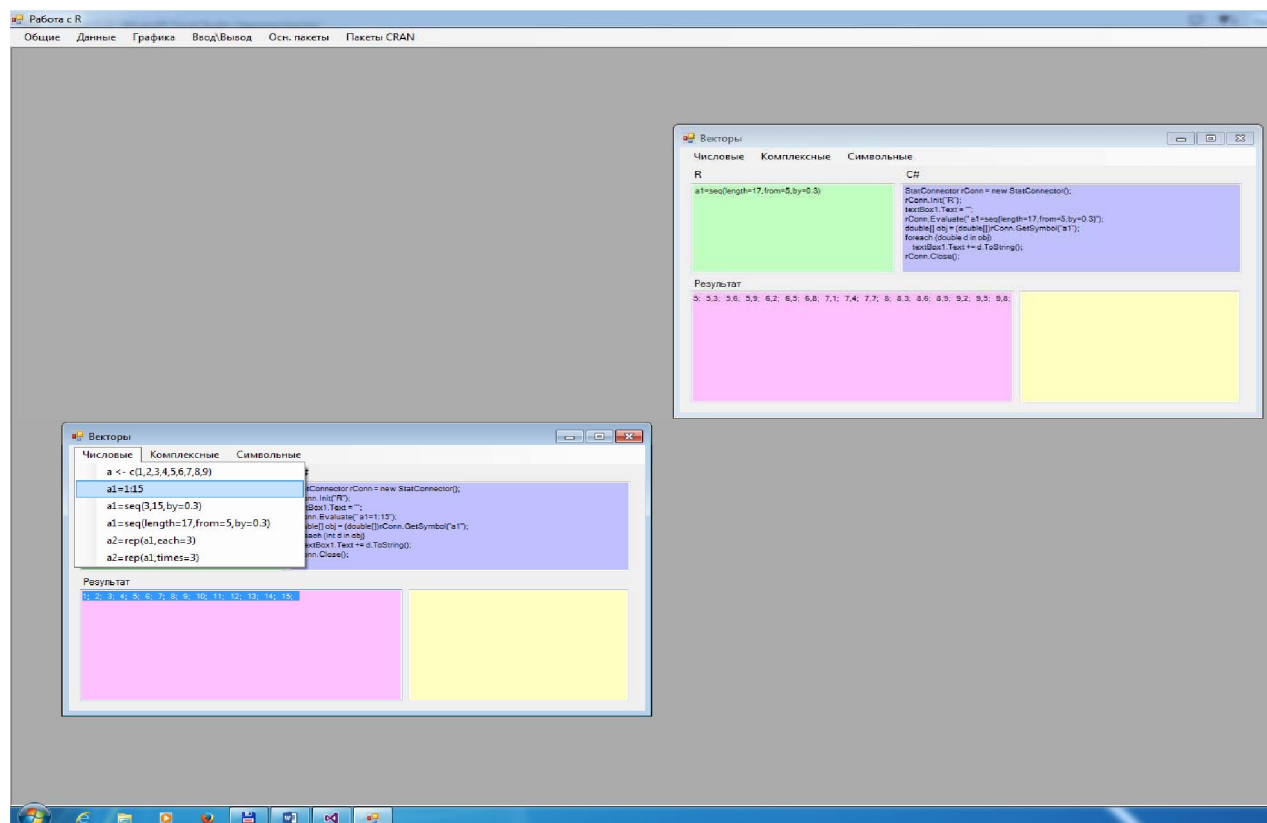


Рисунок 1 – Внешний вид тестовой программы

Часть программного кода тестовой программы.

```
using System;
using System.Collections.Generic;
using System.ComponentModel;
using System.Data;
using System.Drawing;
using System.Linq;
using System.Text;
using System.Threading.Tasks;
using System.Windows.Forms;
using StatConnectorCommonLib;
using STATCONNECTORSRVLib;
namespace Billbak_R_C_3
{
    public partial class Danyye : Form
    {
        public Danyye()
        {
            InitializeComponent();
        }
        private void MDan_1_1_Click(object sender, EventArgs e)
        {
            textBox2.Text = " a <- c(1,2,3,4,5,6,7,8,9)";
            textBox3.Text = " StatConnector rConn = new StatConnector(); \r\n rConn.Init(\"R\"); \r\n
            textBox1.Text = \"\"; \r\n rConn.Evaluate(' a <- c(1,2,3,4,5,6,7,8,9)\");";
            textBox3.Text += "\r\n double[] obj = (double[])rConn.GetSymbol(\"a\"); \r\n foreach (double d in
            obj) \r\n     textBox1.Text += d.ToString(); \r\n rConn.Close();";
            StatConnector rConn = new StatConnector();
            rConn.Init("R");
            textBox1.Text = "";
        }
    }
}
```

```

rConn.Evaluate(" a <- c(1,2,3,4,5,6,7,8,9)");
double[] obj = (double[])rConn.GetSymbol("a");
foreach (double d in obj)
    textBox1.Text += d.ToString() + "; ";
rConn.Close();
}

```

Заклучение. Тестовый пример, рассмотренный в данной статье, показывает большие перспективы использования интеграции среды R в среду C# при построении программных комплексов в данной среде. К минусам данного подхода можно отнести достаточно сложный способ интеграции среды R, а также недостаточно устойчивую работу программы. Однако существуют и другие способы интеграции среды R в среду C#, которые будут рассмотрены позднее в других статьях.

ЛИТЕРАТУРА

- [1] Venables W.N., Smith D.M., R Core Team, An Introduction to R, Notes on R: A Programming Environment for Data Analysis and Graphics, 2015, Version 3.2.3 (2015-12-10), 99 p.
- [2] R Core Team, R Installation and Administration, 2015, Version 3.2.3 (2015-12-10), 73 p.
- [3] R Core Team, R Data Import/Export, 2015, Version 3.2.3 (2015-12-10), 33 p.
- [4] R Core Team, Writing R Extensions, 2015, Version 3.2.3 (2015-12-10), 167 p.
- [5] Гайдьшев И. Анализ и обработка данных. Специальный справочник. - СПб: Питер. - 2001, 752 с.
- [6] Кабаков Р.И. R в действии. Анализ и визуализация данных на языке R. - ДМК Пресс. - 2014, 580 с.
- [7] Троелсен Э. Язык программирования C# 5.0 и платформа .NET 4.5, 6-е изд. - ООО «И.Д. Вильямс». - 2011. - 1392 с.

REFERENCES

- [1] Venables W.N., Smith D.M., R Core Team, An Introduction to R, Notes on R: A Programming Environment for Data Analysis and Graphics, **2015**, Version 3.2.3 (2015-12-10), 99 p (in Eng.).
- [2] R Core Team, R Installation and Administration, 2015, Version 3.2.3 (2015-12-10), 73 p (in Eng.).
- [3] R Core Team, R Data Import/Export, **2015**, Version 3.2.3 (2015-12-10), 33 p (in Eng.).
- [4] R Core Team, Writing R Extensions, **2015**, Version 3.2.3 (2015-12-10), 167 p (in Eng.).
- [5] Gaydyshev I. Analysis and processing of data. Special reference. St. Petersburg: Peter. **2001**, 752 p. (in Rus.).
- [6] Kabakov R.I. R in action. Analysis and visualization of data on R. language - DMK Press. **2014**, 580 p. (in Rus.).
- [7] Troelsen E. C# 5.0 programming language and platform .NET 4.5, 6 th ed. ООО "ID Williams. " **2011**. 1392 p. (in Rus.).

БАҒДАРЛАМАЛЫҚ R ОРТАНЫҢ C# ОРТАСЫНА БІРІКТІРІЛУІ

Н. Литвиненко

ҚР БҒМ ҒК Математика және математикалық модельдеу институты, Алматы, Қазақстан

Түйін сөздер: R ортасы, математикалық модель (үлгі), жасап шығару ортасы, ортаның біріктірілуі, rscrproхu, rcom, қолданбалы бағдарламалар топтамасы (ҚБТ).

Аннотация. Ғылымның және өндірістің әртүрлі салаларында қолданбалы есептерді компьютерлік модельдеген кезде туындайтын негізгі бағыттардың бірі компьютерлік есептеу орталарын біріктіру болып табылады. Бұл проблема жаңа емес, алайда оның өзектілігі уақыт өткен сайын үнемі артып отыр. MATLAB, MAPLE, R, SPSS секілді ең кең таралған қолданбалы бағдарламалардың топтамалары тамаша, жақсы бапталған, ретке келтірілген, жылдам жұмыс істейтін кітапханалардың және әртүрлі есептеу орталарында математикалық есептерді шешкен, мәліметтерді статистикалық өңдеген, мәліметтердің үлкен көлемдерін өңдеген және т.с. кезде ең әр алуан бағыттарды қамтитын басқа құрамдас бөліктердің орасан зор санына ие. Жасалынып жатқан жобалардың бір бөлігі болып табылатын бағдарламалық модульдерді жасап шығарған кезде осы топтамалардың ішінен дайын да болып отырған құрамдас бөліктерді қолданған мақсатқа лайықты болар-ақ еді. Алайда, шеттегі топтамаларды пайдалану міндеті әртүрлі себептерге орай айтарлықтай күрделі болып табылады. Бұл мақалада R топтамасының жекелеген модульдерін C# ортасына біріктірудің негізгі қағидалары сипатталып отыр. R ҚБТ мәліметтерді статистикалық өңдеуге бағдар алған, еркін таралатын бағдарламалық өнім болып табылады. Қазіргі таңда R ҚБТ құрамында статистикадағы ең әр алуан бағыттарды қамтитын 7700 аса топтама бар.

Бұл амал-тәсіл «Әскери ұжымға әлеуметтік-психологиялық диагноз қоюдың автоматтандырылған әдістемесі» гранттық жобасын жасаған кезде қолданылатын болады.

Поступила 13.03.2016 г.