

Экспертная оценка программных продуктов для импутирования пропущенных значений в массивах данных

*Глазунов Артем Дмитриевич
Волжский политехнический институт
студент*

*Рыбанов Александр Александрович
Волжский политехнический институт
Кандидат технических наук, доцент, заведующий кафедрой «Информатика и технология программирования»*

Аннотация

В статье рассмотрены понятия и представлены результаты сравнительного анализа по методу Саати программных продуктов для восстановления пропусков в массиве данных.

Ключевые слова: пропуски в массиве данных, метод Саати.

Expert evaluation of software products for imputation of missed values in data materials

*Glazunov Artem Dmitrievich
Volzhsky Polytechnical Institute
student*

*Rybanov Aleksandr Aleksandrovich
Volzhsky Polytechnical Institute
Ph.D., Associate Professor, Head of the Department "Computer technology and programming"*

Abstract

The article deals with concepts and presents the results of a comparative analysis using the Saati method of software products for recovering omissions in an array of data.

Keywords: Saati method, omissions in the data array.

Причинами появления пропусков в массивах данных являются невозможность получения, сокрытие информации или её искажение. Для импутирования пропусков в массиве данных используются различные программные средства, позволяющие производить восстановление их различными алгоритмами.

В данной статье проведен анализ программных продуктов (MS Excel, DeductorStudio, SAS/ETS Software, SPSS Statistics, Stata) позволяющих

восстанавливать пропущенные значения, и определены их весовые коэффициенты показателей качества методом Саати [2,6].

В качестве сравнительных критериев программного обеспечения были выбраны определенные показатели:

1. А1 - Набор алгоритмов восстановления
2. А2 - Формат входных данных
3. А3 - Скорость восстановления данных
4. А4 - Точность восстановления данных
5. А5 - Простота работы с ПО

Используя аналитическую процедуру Саати, определим вес каждого критерия качества [2].

Таблица 1. Значения коэффициентов матрицы парных сравнений

X_{ij}	Значение
1	i -ый и j -ый критерий примерно равноценны
3	i -ый критерий немного предпочтительнее j -го
5	i -ый критерий предпочтительнее j -го
7	i -ый критерий значительно предпочтительнее j -го
9	i -ый критерий явно предпочтительнее j -го

Таблица 2. Матрица парных сравнений

	А1	А2	А3	А4	А5	Среднее геометрическое	Веса критериев
А1	1	3	5	7	9	3,94	0,50
А2	1/3	1	5	5	7	2,26	0,29
А3	1/5	1/5	1	3	5	0,90	0,11
А4	1/7	1/5	1/3	1	7	0,58	0,07
А5	1/9	1/7	1/5	1/7	1	0,21	0,03
Сумма						7,89	1,00

Диаграмма весовых коэффициентов для показателей А1, А2, А3, А4, А5 показана на рисунке 1.

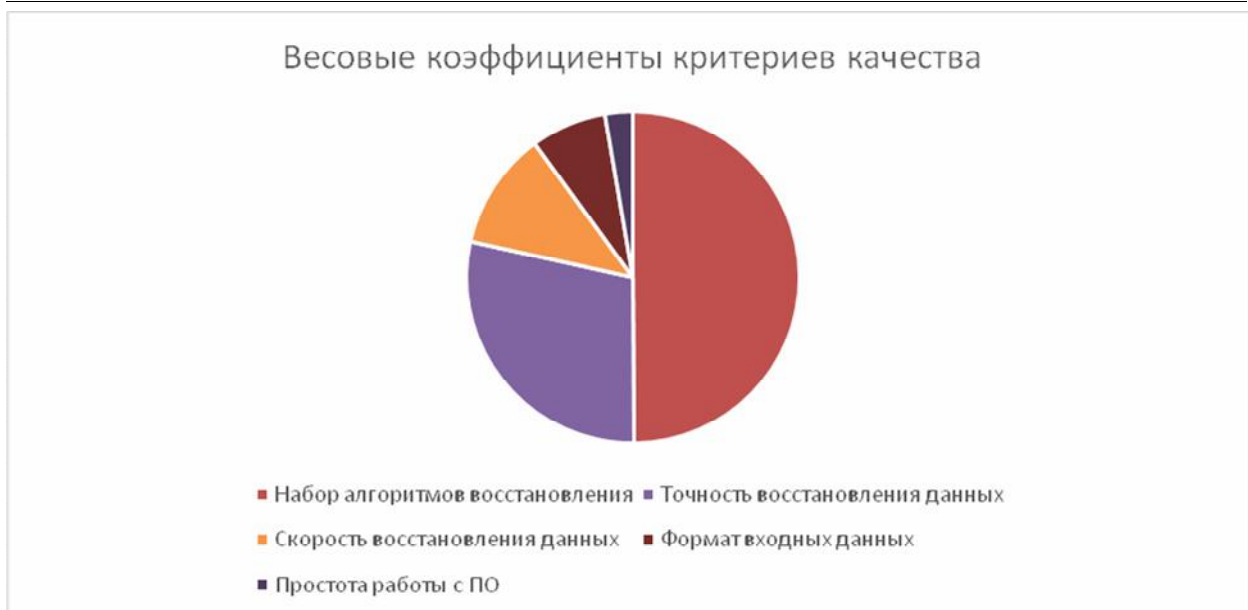


Рисунок 1. Коэффициенты показателей качества

Проверим матрицу сравнений на непротиворечивость.

$$R_1=1,79; R_2= 4,54; R_3= 11,53; R_4= 16,14; R_5=29.$$

Рассчитав величину L , просуммировав произведения сумм столбцов матрицы и весовых коэффициенты:

$$L = 5,48. \text{ ИС} = (L-N)/(N-1) = 0,12.$$

Значение случайной согласованности равно 1,12. $OC = \text{ИС} / \text{СлС} = 0,11$. меньше 0,2, исходя из этого результата производить уточнение матрицы не требуется.

Получив коэффициенты показателей, определим интегральный показатель качества для выбранных программных продуктов, позволяющих восстанавливать пропуски в данных:

1. Deductor Studio;
2. SPSS Statistics;
3. MS Excel;
4. Stata;
5. SAS/ETS Software.

Значения категориальной шкалы составляют от 0 до 7 (где 0 – качество не удовлетворительно, 7 – наилучший показатель) для функциональных решений выбранных программ [4,5].

Значения весовых коэффициентов a_i , соответствующие возможностям программных средств:

1. Набор алгоритмов восстановления: $a_1 = 0,50$;
2. Точность восстановления данных: $a_2 = 0,29$;
3. Скорость восстановления данных: $a_3 = 0,11$;
4. Формат входных данных: $a_4 = 0,07$;
5. Простота работы с ПО: $a_5 = 0,03$;

$$\text{где } \sum a_i = 1.$$

Выбрав шкалу, определим количественные значения функциональных возможностей X_{ij} (таблица 3), а также, рассчитаем показатели качества для тестируемых программных средств.

Таблица 3. Показатели качества критериев

Критерии	Весовые коэффициенты	Программные продукты					Базовые значения
		Deductor Studio	SPSS Statistics	MS Excel	Stata	SAS/ETS Software	
Набор алгоритмов восстановления	0,50	7	7	2	4	4	5,00
Точность восстановления данных	0,29	5	7	5	5	4	4,80
Скорость восстановления данных	0,11	7	6	5	3	6	5,20
Формат входных данных	0,07	6	5	3	5	4	4,60
Простота работы с ПО	0,03	5	6	7	5	3	4,80
Интегральные показатель качества Q		6,47	6,58	2,13	4,57	4,90	4,93

Значения характеристик функциональных возможностей показаны на рисунке 2.

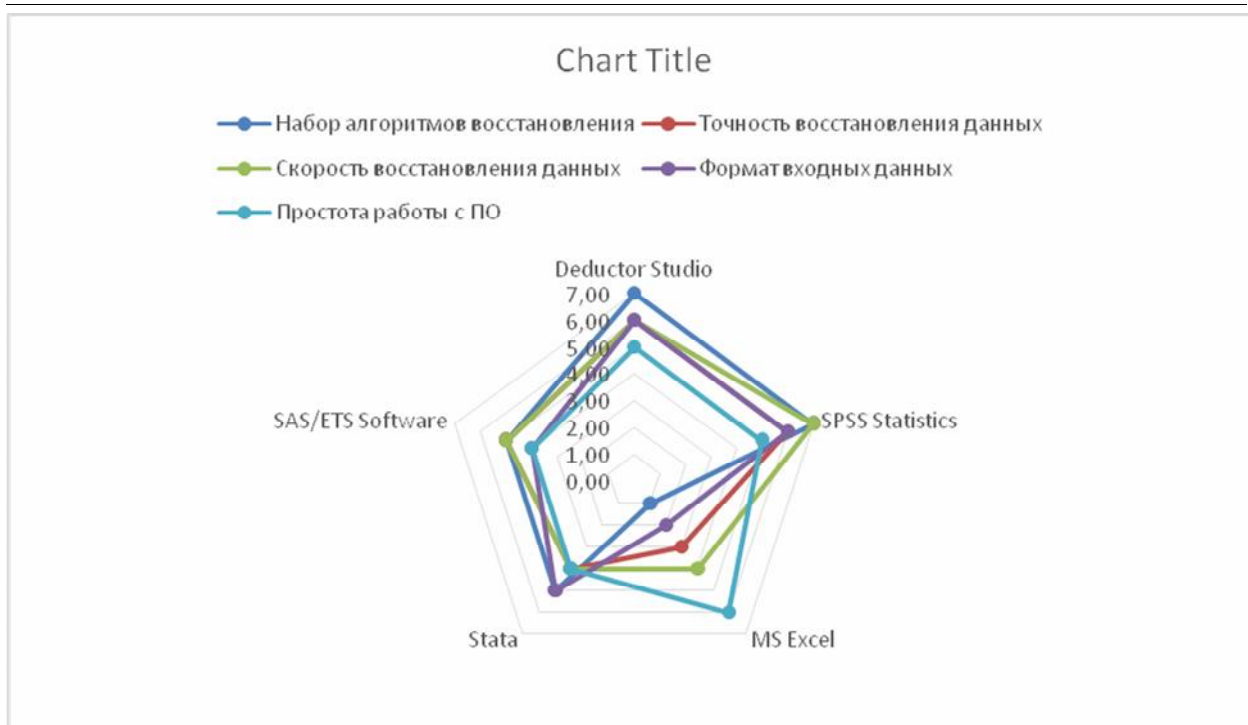


Рисунок 2. Лепестковая диаграмма

Сравнительный анализ ПО для восстановления пропусков в массиве данных показал, что из всех тестируемых программных систем, только Deductor Studio и SPSS Statistics имеют значение интегрального показателя качества, превышающий базовое значение, у других же программных аналогов оно оказалось ниже базовой отметки. Так же анализ показал, что Deductor Studio и SPSS Statistics надо повысить показатели по таким критериям как: «Точность восстановления данных» и «Формат входных данных».

Оценочная методика программных продуктов позволила количественно оценить качество эффективности реализованных функций и выявила те функции, которые реализованы в недостаточной степени некоторых протестированных программных продуктах.

Библиографический список

1. Богушенков А.С., Рыбанов А.А. Разработка и исследование алгоритмов автоматизированной системы учета и поиска информации по пакетам труб на основе технологии QR-кода // Молодой ученый. 2015. № 4 (84). С. 47-52.
2. Кондрацкий Д.Е., Рыбанов А.А. Исследование методов и алгоритмов автоматизированной системы оценки альтернативных вариантов методом Т.Саати // NovaInfo.Ru. 2016. Т. 3. № 46. С. 107-116.
3. Рыбанов А. Определение весовых коэффициентов сложности тем учебного курса на основе алгоритма Саати // Педагогические измерения. 2014. № 4. С. 21-28.
4. Рыбанов А.А., Макушкина Л.А. Технология определения весовых

- коэффициентов сложности тем дистанционного курса на основе алгоритма Саати // Открытое и дистанционное образование. 2016. № 1 (61). С. 69-79.
5. Рыбанов А.А., Усмонов М.С.О., Попов Ф.А., Ануфриева Н.Ю., Бубарева О.А. Информационные системы и технологии // Научный редактор: И. А. Рудакова/ Центр научной мысли (г. Таганрог). Москва, 2013. Том Часть 4 Информационные системы и технологии.
 6. Морозов А.О., Рыбанов А.А. Экспертная оценка программных продуктов для расчета метрических характеристик физической схемы базы данных // Современные научные исследования и инновации. 2015. № 1-1 (45). С. 97-102.
 7. Rybanov A.A. Set of criteria for efficiency of the process forming the answers to multiple-choice test items. Turkish Online Journal of Distance Education. 2013. T.14. №. 1. С. 75-84