

ПРИМЕНЕНИЕ РАЗВЕДОЧНОГО АНАЛИЗА ДАННЫХ ДЛЯ КЛАСТЕРИЗАЦИИ СТРУКТУР РОБОТОТЕХНИЧЕСКИХ СБОРОЧНЫХ КОМПЛЕКСОВ

Яковлев Д.Д.¹, Петров Д.Ю.²

¹*Институт проблем точной механики и управления Российской академии наук,
Саратов;*

²*Институт прикладной математики и механики, Донецк*

Ключевые слова: дендрограмма, разведочный анализ, иерархическая кластеризация, Loginom, NCSS.

Аннотация. Проведена классификация робототехнических сборочных комплексов. Определены основные элементы и параметры, используемые для сравнения робототехнических сборочных комплексов. Показана структура разведочного анализа данных в программном комплексе Loginom. В результате проведения разведочного анализа были получены исходные данные для кластеризации базовых структур робототехнических сборочных комплексов. Проведено сравнение результатов кластеризации робототехнических сборочных комплексов методами Уорда и гибкой стратегии в программном комплексе NCSS. На основе результатов кластерного анализа определяются основные типы базовых структур РСК.

APPLICATION OF EXPLORATORY DATA ANALYSIS FOR CLUSTERIZATION OF ROBOTIC ASSEMBLY COMPLEXES STRUCTURES

Yakovlev D.D.¹, Petrov D.Yu.²

¹*Institute of Precision Mechanics and Control of the Russian Academy of Science,
Saratov;*

²*Institute of Applied Mathematics and Mechanics, Donetsk*

Keywords: dendrogram, exploratory analysis, hierarchical clustering, Loginom, NCSS.

Abstract. The classification of robotic assembly complexes is carried out. The main elements and indicators used for comparison of robotic assembly complexes are defined. The structure of exploratory data analysis in Loginom software package is shown. As a result of exploratory analysis, the initial data for clustering of basic structures of robotic assembly complexes were obtained. The results of clustering of robotic assembly complexes by Ward and flexible strategy methods in NCSS software package were compared. On the basis of the results of cluster analysis the main types of basic structures of robotic assembly complexes are determined.

Развитие отечественного машиностроения сдерживается отсутствием необходимого количества квалифицированных кадров. Проблему дефицита кадров во всех отраслях можно решить при помощи роботизации [1, 2]. Современные робототехнические сборочные комплексы (РСК) используются для выполнения сложных задач в различных отраслях промышленности. Они интегрируют промышленных роботов, средства автоматизации, системы контроля и управления с использованием искусственного интеллекта. Использование РСК в машиностроении способствует увеличению производственной эффективности, сокращению времени обработки изделий, минимизации цикловых заделов и повышению производительности труда [3].

РСК обладают высокой стоимостью, и не оптимальность их структур оказывает существенное влияние на сроки их окупаемости [4]. Анализ структур базируется на многомерной классификации РСК. Она проводится по различным критериям, таким как характеристики изделия, тип используемых роботов, уровень автоматизации, структура управления и др. Структуры современных РСК включают более 30 типов элементов [5]. К основным элементам РСК относятся:

- 1) роботы-манипуляторы;
- 2) средства транспортировки деталей (конвейеры, мобильные роботы);
- 3) рабочие станции, системы подачи материалов;
- 4) системы управления (центральные контроллеры, распределённые управляющие устройства, системы распределенного ввода-вывода);
- 5) датчики (для обратной связи, контроля качества, измерений);
- 6) программное обеспечение (системы управления производственными процессами, интеграция с CAD/CAM);
- 7) интерфейсы взаимодействия с оператором (HMI – Human Machine Interface).

Такая детализированная классификация позволяет более точно создавать кластеры при использовании методов иерархической кластеризации.

Результатом иерархической кластеризации служит дендрограмма – это графическое представление данных, которое используется для визуализации структуры данных. В данной работе кластеризация структур РСК будет проводиться методами Уорда и гибкой стратегии в программном комплексе NCSS [6, 7].

Кластеризация РСК предполагает предварительное проведение разведочного анализа данных с целью выявления наиболее общих зависимостей, закономерностей и тенденций, характера и свойств анализируемых данных, законов распределения анализируемых величин. Он необходим для нахождения связей между переменными в ситуациях, когда отсутствуют (или недостаточны) априорные представления о природе этих связей. В разведочном анализе учитывается и сравнивается большое число параметров объекта исследования, а для поиска закономерностей используются методы: процедура анализа распределений переменных, корреляционный анализ с целью поиска коэффициентов, превосходящих по величине определенные пороговые значения, факторный анализ, дискриминантный анализ, многомерное шкалирование, визуальный анализ гистограмм и т.д.

При сравнении РСК используются параметры: стоимость РСК, руб.; затраты на эксплуатацию, количество технологических операций, шт.; масса, кг; максимальная нагрузка, кг; габариты РСК (длина, ширина, высота), мм; габариты изделия (длина, ширина, высота), мм; точность позиционирования, мм; расход электроэнергии, кВт/ч, производительность шт./ч и др. [8]. Многомерный анализ этих данных позволяет более точно создавать кластеры при использовании методов иерархической кластеризации. Для разведочного анализа подготовлены данные по 97 существующим РСК (рис. 1).

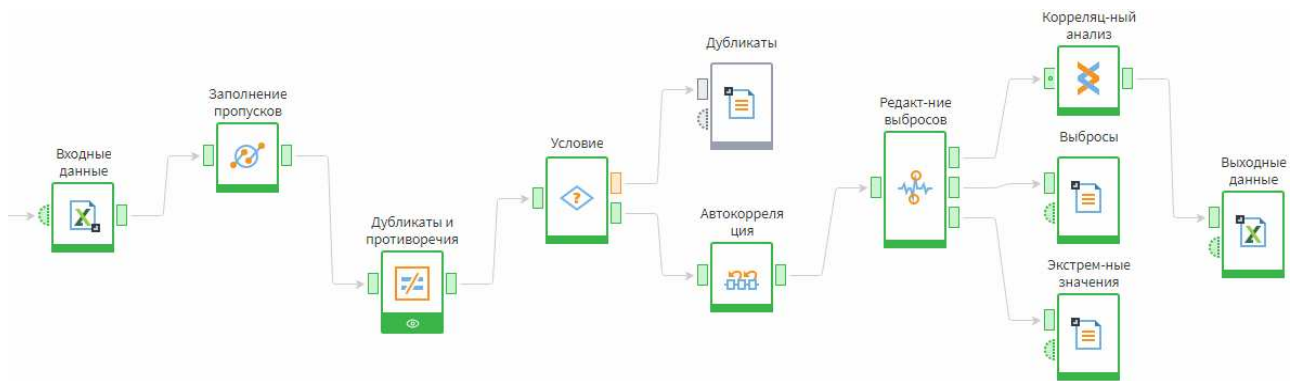


Рис. 1. Структура разведочного анализа в программном комплексе Loginom

Разведочный анализ в программном комплексе Loginom [9], состоит из следующих шагов.

1) На вход подаются данные, представленные в виде таблицы в Excel файле.

2) Блок «Заполнение пропусков» – предназначен для автоматического заполнения пропущенных значений в наборах данных. Для каждого столбца исходного набора данных выбирается наиболее подходящий метод заполнения пропусков.

3) Блок «Дубликаты и противоречия» – позволяет выявить в исходной выборке данных дублирующие и противоречивые записи. К дубликатам относятся записи в таблице, все входные и выходные поля которых одинаковые. Дубликаты приводят к избыточности, увеличивают объем выборки, при этом не повышая информативность данных. К противоречиям относятся записи в таблице, у которых все входные поля одинаковые, но отличаются хотя бы по одному выходному полю. Противоречия приводят к искажению результата анализа и снижают качество моделей, поскольку нарушают общие закономерности в данных, обнаружение которых и является целью исследования.

4) Если в блоке «Дубликаты и противоречия» были некорректные данные, блок «Условие» не даёт им попасть в последующие блоки и записывает такие данные в отдельный текстовый файл.

5) Блок «Автокорреляция» – позволяет вычислять автокорреляцию временных рядов, количество отсчетов задаётся лицом, принимающим решение (ЛПР).

6) Блок «Редактирование выбросов» – предназначен для автоматической корректировки выбросов и экстремальных значений в наборах данных. Для каждого поля набора данных критерии определения выбросов и экстремальных значений задаются ЛПР с помощью указания допустимого стандартного отклонения или интерквартильного размаха. Под выбросами при этом подразумеваются значения данных, существенно отклоняющиеся от средних. Под экстремальными подразумеваются такие значения, которые настолько сильно отклоняются от типичных значений, что перестают соответствовать логике исследуемых процессов и явлений. Выбросы и экстремальные значения записываются в отдельные текстовые файлы.

7) Блок «Корреляционный анализ» – выявляет, на основании вычисляемых коэффициентов корреляции, взаимосвязь между рядами данных входного набора.

Применяется для оценки предполагаемой зависимости факторов. ЛПР может выбрать коэффициенты для оценки корреляции: коэффициент корреляции Пирсона, коэффициент Тау-в Кендалла, экстремум взаимнокорреляционной функции или коэффициент корреляции Спирмена.

В результате проведения разведочного анализа в программном комплексе Loginom, данные были объединены в 10 групп. Распределение РСК по группам: 1) 10 шт.; 2) 9 шт.; 3) 12 шт.; 4) 9 шт.; 5) 9 шт.; 6) 11 шт.; 7) 10 шт.; 8) 9 шт.; 9) 10 шт.; 10) 8 шт. Эти группы являются исходными данными для кластеризации базовых структур РСК.

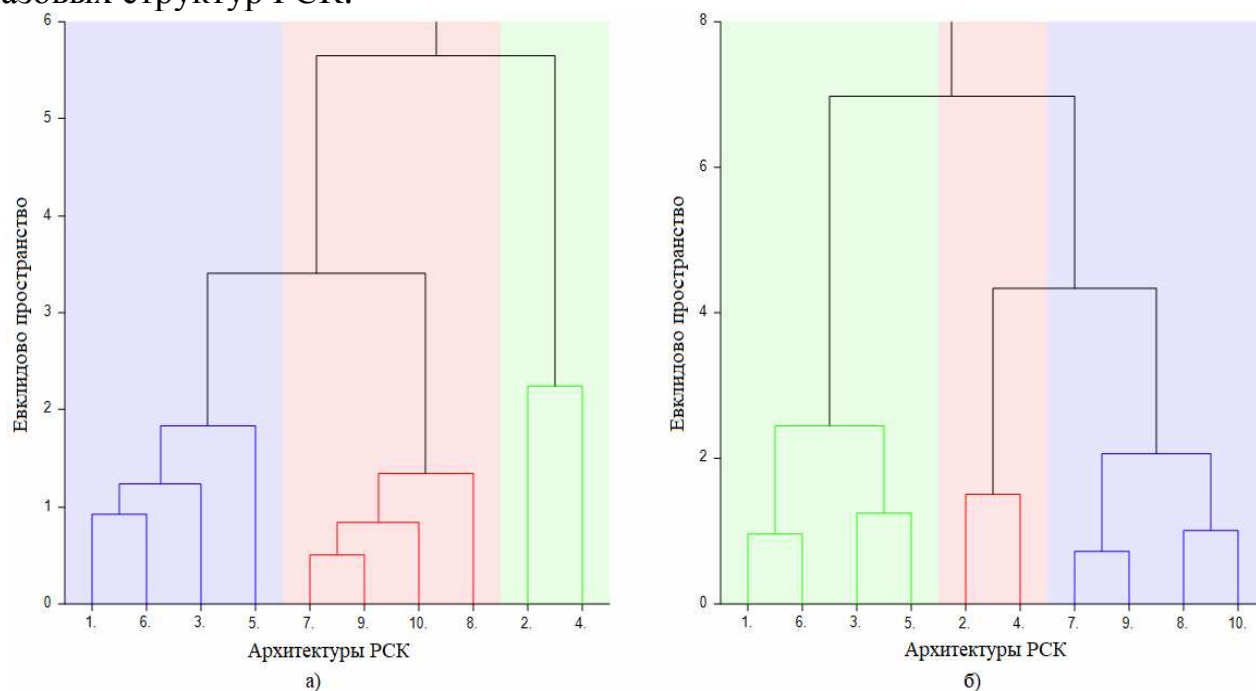


Рис. 2. Иерархическая кластеризация структур РСК: а) методом Уорда, б) методом гибкой стратегии

Результат иерархической кластеризации показан на рисунке 2. Из полученных дендрограмм можно сделать вывод, что набор РСК можно представить в виде 3 кластеров, отмеченных разными цветами (красным, синим, зелёным), а именно: первый кластер – РСК под номерами 1-3-5-6; второй кластер – РСК под номерами 7-8-9-10; третий кластер – РСК под номерами 2-4. На основе результатов кластерного анализа определяются основные типы базовых структур РСК.

Список литературы

1. Орешкин М.С. Роботизация – это прямой путь к преодолению дефицита кадров, который касается всех отраслей [Электронный ресурс] // Ведомости. – Режим доступа: <https://www.vedomosti.ru/technology/news/2024/10/03/1066280-robotizatsiei>.
2. Романова Е.Е., Пелымская И.С. Анализ и оценка влияния роботизации на уровень безработицы // Весенние дни науки ИнЭУ: Сборник докладов международной конференции студентов и молодых ученых. 17-20 апреля 2024г. Электронное научное издание; ФГАОУ ВО «УрФУ имени первого Президента России Б.Н. Ельцина». – Екатеринбург: Издательский Дом «Ажур», 2024. – С. 345-350.
3. Староватова Д.А. Связь уровня роботизации и производительности труда: важен ли масштаб бизнеса? // Journal of new economy. – 2023. – Т. 24, №1. – С. 81-103. DOI: 10.29141/2658-5081-2023-24-1-4.

4. Яковлев Д.Д., Петров Д.Ю., Большаков А.А. Критерии оптимизации архитектур робототехнических сборочных комплексов // Математические методы в технологиях и технике. – 2024. – № 8. – С. 28-31.
5. Хомченко В.Г. Робототехнические системы. – Омск: Изд-во ОмГТУ, 2016. – 195 с.
6. NCSS: Data Analysis & Graphic [Электронный ресурс] // Statistical Software. – Режим доступа: <https://www.ncss.com/software/ncss/>
7. Lance G.N., Williams W.T. A General Theory of Classificatory Sorting Strategies: 1. Hierarchical Systems // Computer Journal. 1967, vol. 9, pp. 373-380.
8. Jiang Y., Huang Z., Yang B., Yang W. A review of robotic assembly strategies for the full operation procedure: planning, execution and evaluation // Robotics and Computer-Integrated Manufacturing. 2022, vol. 78, p. 102366.
9. Аналитическая платформа Loginom [Электронный ресурс] // О платформе Loginom. – Режим доступа: <https://loginom.ru/platform>.

Сведения об авторах:

Яковлев Денис Дмитриевич – аспирант;

Петров Дмитрий Юрьевич – к.т.н., доцент, старший научный сотрудник.