

Байшев Анатолий Викторович, аспирант кафедры «Цифровых технологий и дизайна», Хакасский государственный университет им. Н.Ф. Катанова

ПЕРСПЕКТИВЫ ОБРАБОТКИ ДЕФЕКТОВ В ДАННЫХ ОДНОРОДНЫХ ТЕХНИЧЕСКИХ СИСТЕМ

Аннотация: В статье рассмотрены перспективы обработки дефектных данных технических систем на основании информации о данных других однородных ей систем, с использованием установленных относительно таковых закономерностей. Приведена методика вменения дефектных данных однородных технических систем и осуществлено ее сравнение с методами квадратичной и линейной интерполяции используемых во многих программных комплексах диспетчерского управления и сбора данных для заполнения отсутствующих данных.

Ключевые слова: системный анализ, управление, однородные системы, выбросы, пропуски данных, интерполяция данных.

Abstract: The article considers the prospects for processing defective data of technical systems on the basis of information about the data of other systems homogeneous to it, using the regularities established with respect to such. A technique for imputing defective data of homogeneous technical systems is presented and compared with the methods of quadratic and linear interpolation used in many software systems for dispatch control and data collection to fill in missing data.

Key words: system analysis, control, homogeneous systems, outliers, data gaps, data interpolation.

Введение. Большую проблему в вопросе работы с данными различных технических систем являет собой наличие в них различных дефектов, в

частности пропусков и выбросов своим появлением обусловленных влиянием различных, в том числе, случайных факторов, таких как, например, отказ регистрирующей техники по какой-либо причине [1 - 4]. Обработка и устранение таких, с целью установления приближенного к реальности отображения того или иного технического процесса является собой весьма сложную и актуальную задачу [5 - 6]. Особенно это актуально в части управления техническими системами, где может быть особенно полезно построение некоей модели данных (желательно максимально приближенной к реальным нормальным данным) о том или ином параметре системы, отражающей приблизительные показатели нормального течения процесса в момент наличия в них выбросов или отсутствующих данных, как это изображено на рисунке 1.

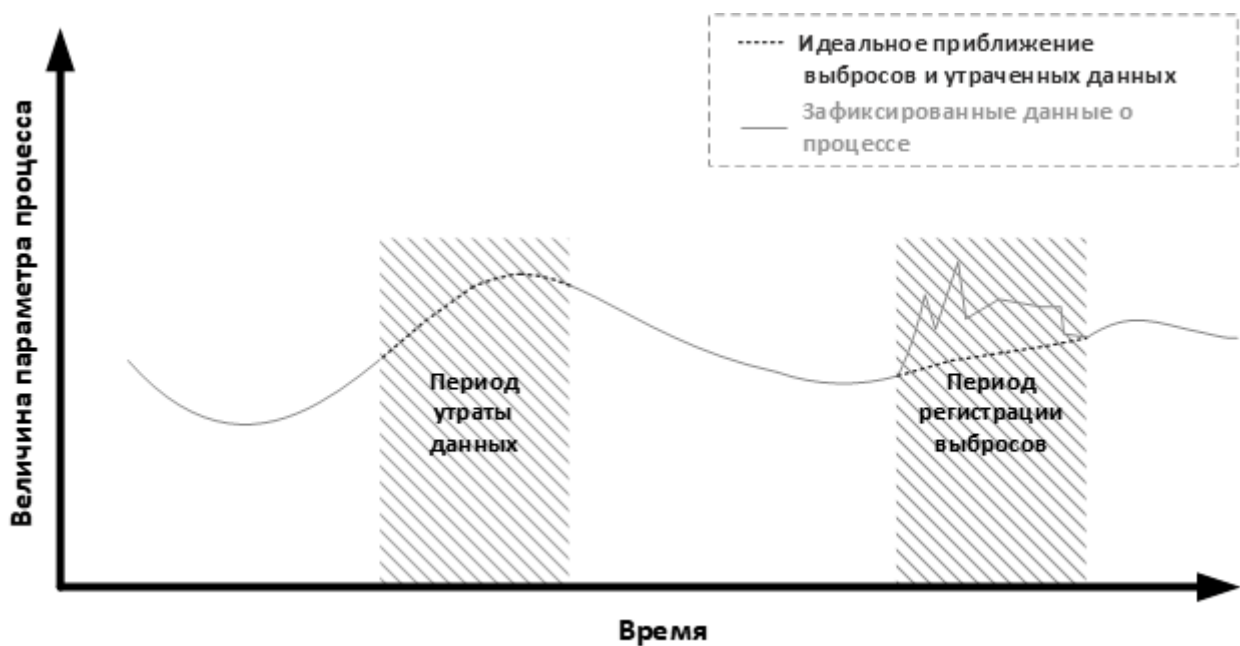


Рисунок 1 – Зарегистрированные данные о процессе с дефектами и пропусками, а также идеальные(реальные) показатели процесса

В свою очередь, различные типы систем могут обладать свойствами, полезными для формирования приблизительных показателей нормального течения технических процессов (модели данных) в моменты регистрации выбросов или отсутствующих значений. На этом фоне особо актуально использование системного анализа для проведения исследований и определения

таких свойств для различных видов систем.

В работе сосредоточено внимание на однородных технических системах (далее ОТС), то есть технических системах одинаковых по конструкции и выполняющих одну и ту же полезную функцию [7]. Эти системы весьма распространены в современном мире и их популярность можно объяснить, в том числе, экономическими предпосылками ведь, как правило, применение нескольких однородных систем выгоднее применения лишь одной большой таковой, так как дает более широкие возможности для оптимизации объемов хранения запасных частей, позволяет гибко настраивать эксплуатационные процессы, планировать ремонтные программы и т.д.

ОТС широко применяют в самых разных сферах, так в транспортной сфере в качестве примера таковых можно обозначить: железнодорожные вагоны, вагоны метро, автомобили одной марки модели и комплектации. Среди оборудования в области электроэнергетики к ним можно отнести некоторые множества трансформаторов, генераторов, опор линий электропередачи и т.д. В промышленной сфере получило широкое распространение применение нескольких однородных производственных линий в рамках одного предприятия.

Целью работы является определение перспективности обработки дефектных данных технической системы на основании информации о данных других однородных ей систем с использованием установленных относительно таких ОТС закономерностей. Актуальность исследования заключается в том, что результаты работы могут быть полезны при работе с данными различных технических систем, в том числе, для включения их в алгоритмы программных пакетов диспетчерского управления и сбора данных – англ. Supervisory Control And Data Acquisition (SCADA).

Материалы и методы исследования. Для того, чтобы осуществлять обработку дефектных данных, на начальном этапе исследования необходимо определить подход к обоснованию однородности тех или иных технических систем. Этот ответственный этап крайне важен и на может быть реализован путем анализа нормативной документации, инструкций и т.п., которые для

предполагаемых множеств ОТС должны быть одинаковы. Для реализации этого процесса возможно также привлечение отраслевых экспертов, обладающих знаниями о предполагаемых к анализу технических системах.

Выявив множества ОТС, можно детальнее исследовать данные того или иного из них. В данных можно с помощью различных методов обнаружить дефекты в виде пропусков и выбросов. Далее можно попытаться исправить эти дефекты основываясь на свойствах исследуемых систем и информации о режиме их работы.

В части анализа ОТС представляет особый интерес сравнение значений их параметров при практически равных условиях внешнего воздействия на них, так как при таких они должны быть практически одинаковыми. Это говорит о том, что дефектные данные той или иной системы можно попытаться интерполировать с помощью данных других однородных ей систем, если известна информация о том, что такие системы находились в условиях практически равного воздействия на них. Особо это актуально для заполнения данными относительно небольших промежутков времени, так как чем длительнее промежутки времени с дефектами данных, тем выше неопределенность относительно нахождения системы в тех или иных условиях на них.

В вопросе попыток интерполяции дефектных данных ОТС их приближенными к реальным значениями может помочь исследование взаимосвязей между данными до появления в них выбросов или пропусков и после.

На рисунке 2 изображены данные о некотором параметре для пяти ОТС, работающих на рассматриваемом временном промежутке в практически равных условиях. Видно, что данные о выбранном параметре систем практически идентичны.

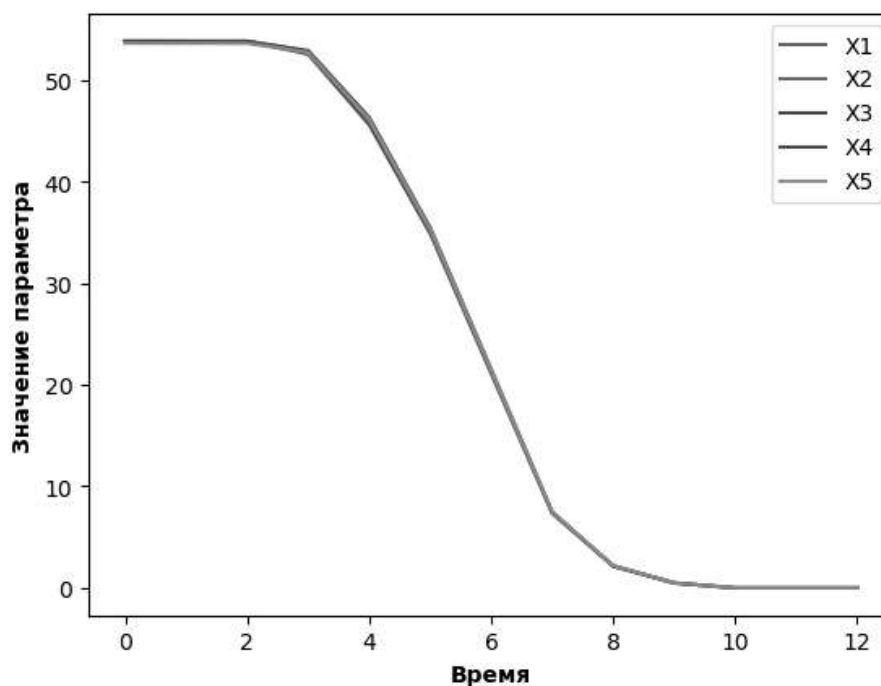


Рисунок 2 – Изменение некоторого параметра ОТС, находящихся в практически равных условиях внешнего воздействия на них

Предположим, что с момента времени «2» по момент времени «4» были утрачены, либо на них регистрировалось отсутствие данных о системе «X2». Это значения выделены шрифтом и цветом в таблице 1.

Таблица 1 – Исследуемые данные ОТС

	X1	X2	X3	X4	X5	Линейная интерполяция «X2»	Квадратичная интерполяция «X2»
0	53,8	53,8	53,8	53,8	53,6	53,8	53,8
1	53,8	53,8	53,8	53,8	53,6	53,8	53,8
2	53,8	53,8	53,8	53,8	53,6	49,2	52,7
3	52,8	52,8	52,8	52,8	52,6	44,6	50,4
4	46,1	46,1	46,1	46,1	45,9	39,9	45,2
5	35,4	35,4	35,4	34,4	34,2	35,4	35,4

В рассматриваемой ситуации, на основании анализа данных до инцидента появления дефектных данных и после можно выявить закономерности относительно взаимосвязей данных систем и на его основании определить

значения параметров при нормальном течении процесса. Исходя из информации, интерпретируемой из таблицы 1 видно, что большая часть систем характеризуются одинаковыми данными, в то время как для одной из систем (система «X5») значения параметра закономерно ниже в каждый момент времени относительно значений остальных систем на 0,2 единицы в каждый из моментов времени.

Кроме того, в таблице 1 и на рисунке 3 представлены данные о результатах применения линейной и квадратичной интерполяции для заполнения отсутствующих значений системы «X2», так как такие способы восстановления информации, особенно метод линейной интерполяции, весьма популярны за счет их простоты и применяются на практике в различных программных пакетах SCADA.

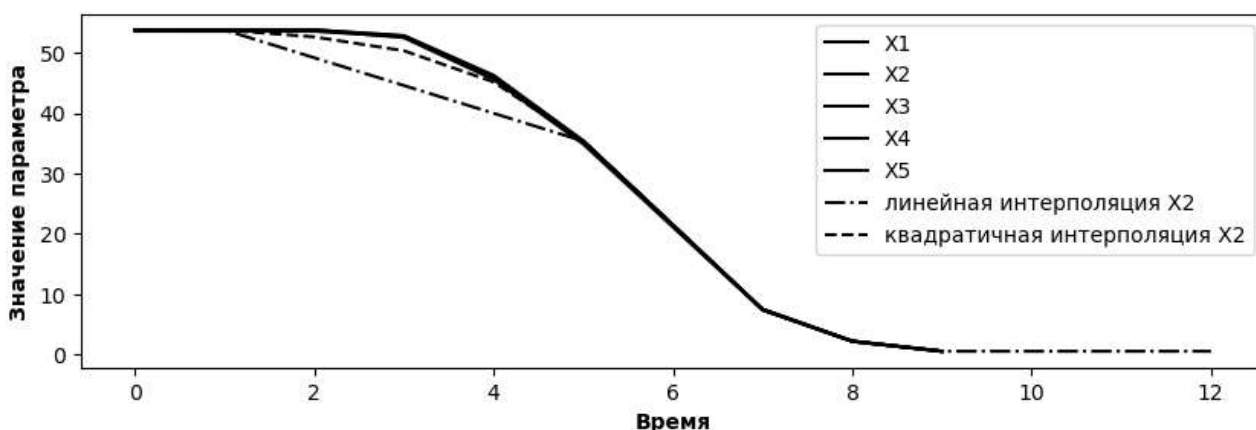


Рисунок 3 – Графики параметра 5 ОТС, а также интерполированные графики для данных системы «X2»

Как следует из рисунка 3, можно сказать о том, что любая из рассмотренных форм интерполяции отсутствующих данных существенно проигрывает данным любой из пяти исследуемых ОТС, находящихся в практически равных условиях внешнего воздействия на них. Однако, на основании анализа данных о параметре систем до и после инцидента с возникновением отсутствия данных о системе «X2», приведенных в таблице 1, можно заметить, что для целей интерполяции данных о «X2» особо хорошо

подходят данные трех систем: «X1», «X3», «X4».

Вывод. Таким образом, видно, что выявление закономерностей относительно взаимосвязей данных ОТС весьма перспективно при решении вопросов определения приблизительных показателей нормального течения процессов в момент наличия в них выбросов или отсутствующих данных, так как может давать большую точность по сравнению с традиционными методами.

Библиографический список:

1. Problems in analyzing time series with gaps and their solution with the winabd software package / Desherevskii, A.V. et al. // *Izvestiya, Atmospheric and Oceanic Physics*. – 2017. – Vol. 53, №7, – P. 659–678 <https://doi.org/10.1134/s0001433817070027>.

2. Ch. Sanjeev Kumar Dash, Ajit Kumar Behera, Satchidananda Dehuri, Ashish Ghosh, An outliers detection and elimination framework in classification task of data mining, *Decision Analytics Journal*, Vol. 6, 2023, 100164, <https://doi.org/10.1016/j.dajour.2023.100164>.

3. Tianyi Zhao, Yue Sun, Zhuyue Chai, Kuishan Li, An outlier management framework for building performance data and its application to the power consumption data of building energy systems in non-residential buildings, *Journal of Building Engineering*, Vol. 65, 2023, 105688, <https://doi.org/10.1016/j.jobbe.2022.105688>.

4. H.S. Jha, A. Khanal, H.M.D. Seikh, W.J. Lee, A comparative study on outlier detection techniques for noisy production data from unconventional shale reservoirs, *Journal of Natural Gas Science and Engineering*, Vol. 105, 2022, 104720, <https://doi.org/10.1016/j.jngse.2022.104720>.

5. Очистка сенсорных данных в интеллектуальных системах управления отоплением зданий / М. Л. Цымблер [и др.] // *Вестник Южно-Уральского гос. ун-та. Сер. Вычислительная математика и информатика*. – 2021. – №3. – С. 16–36. <https://doi.org/10.14529/cmse210302>.

6. Позолотин В. Е., Султанова Е. А. Применение алгоритмов преобразования данных при анализе временных рядов на предмет устранения

выбросов // Программные системы и вычислительные методы, 2019, №2, С. 33-42. <https://doi.org/10.7256/2454-0714.2019.2.28279>.

7. Проворов, А. В. Техническое творчество: учебное пособие для среднего профессионального образования / А. В. Проворов. 2-е изд. Москва: Издательство Юрайт, 2023. 425 с.