

УДК 681.514

Алгоритмизация процесса мониторинга данных на автономных объектах электропитания в условиях прогнозирования потенциальных неисправностей на основе однородного глубокого нейронного ансамбля

Д.Н. Карпухин, В.Л. Бурковский

Воронежский государственный технический университет, Воронеж, Россия

Используемые в настоящее время интеллектуальные системы мониторинга зачастую не имеют функций прогнозирования состояния системы в режиме реального времени. В данной статье исследуется концепция системы мониторинга, основанная на использовании нейронного ансамбля, который может применяться для прогнозирования потенциальных неисправностей, возникающих на автономных объектах электропитания.

Ключевые слова: мониторинг данных, система обнаружения потенциальных неисправностей, автономный объект электропитания, глубокий нейронный ансамбль.

Algorithmization of the data monitoring process at autonomous power supply facilities in conditions of predicting potential failures based on a homogeneous deep neural ensemble

D.N. Karpukhin, V.L. Burkovsky

Voronezh State Technical University, Voronezh, Russia

Currently used intelligent monitoring systems often do not have the functions of predicting the state of the system in real time. This article explores the concept of a monitoring system based on the use of a neural ensemble, which can be used to predict potential malfunctions occurring at autonomous power supply facilities.

Keywords: data monitoring, potential fault detection system, autonomous power supply facility, deep neural ensemble.

Важнейшей задачей при разработке сложных радиотехнических комплексов (СРТК) является обеспечение высокой надежности. Для этого необходимы эффективные интеллектуальные системы мониторинга, которые будут постоянно отслеживать информацию, протекающую в СРТК. Современные промышленные объекты обладают высокими вычислительными ресурсами и памятью. Тем не менее, выявление потенциальных неисправностей среди огромного объема данных по-прежнему является сложной задачей. Незначительные отклонения системы, которые не были обнаружены вовремя, могут позже перерасти в более серьезные проблемы.

Мониторинг данных является ключевым элементом в обеспечении стабильной и эффективной работы любой системы. Он позволяет отслеживать состояние системы, выявлять проблемы и своевременно реагировать на них. Существует два основных подхода мониторинга данных: реактивный и проактивный [1, 2]. Реактивный мониторинг основан на сборе и анализе данных о текущем состоянии системы, что позволяет реагировать на конкретные аварийные ситуации, которые уже произошли. Действия при в данном подходе предпринимаются только после возникновения проблемы. Проактивный мониторинг в отличие от реактивного, предполагает

постоянный анализ данных, выявление потенциальных рисков и разработку мер по их предотвращению. В данном подходе основной упор сделан на предупреждении проблем. Выбор между реактивным и проактивным мониторингом зависит от конкретных потребностей системы. Для систем с высокой критичностью, таких как автономный объект электропитания (АОЭ), рекомендуется использовать проактивный мониторинг. Для систем с меньшими требованиями можно использовать реактивный мониторинг. Важно понимать, что эти два подхода не являются взаимоисключающими. Оптимальным решением может быть использование комбинации обоих подходов для достижения максимальной эффективности и надежности системы.

Обычно при мониторинге данные собираются из потока телеметрии в лог-файлах различных модулей, и вопросы прогнозирования не рассматриваются. Традиционно аномалии в данных обнаруживают вручную, просматривая телеметрию, собранную в файлах журналов. Данный метод диагностики нуждается в обновлении, так как он не в состоянии адекватно реагировать на быстро меняющиеся условия работы системы и не всегда может точно определить нормальное состояние. Это приводит к тому, что некоторые аномалии в данных могут остаться незамеченными, что чревато негативными последствиями. Для решения этой проблемы необходим комплексный подход, включающий в себя несколько этапов:

1) Разработать алгоритм для определения нормального состояния системы. Данный алгоритм может быть разработан путем анализа исторических данных о работе системы, выявления характерных паттернов и создания модели, которая будет учитывать эти паттерны. Разработанная модель должна быть достаточно гибкой, чтобы адаптироваться к изменениям в условиях работы системы.

2) Разработать алгоритм прогнозирования будущего состояния системы. Данный алгоритм необходим для того, чтобы прогнозировать возможные отклонения в системе и своевременно реагировать на них. Существует множество методов прогнозирования, и выбор конкретного зависит от специфики задачи. Статистические методы, основаны на анализе исторических данных. Они давно зарекомендовали себя в различных областях, но по мере роста объема обрабатываемой информации, статистические модели теряют свою актуальность, не справляясь с быстро меняющимися трендами. Преимущество методов машинного обучения в том, что они способны адаптироваться к изменяющимся условиям. Однако для критически важных решений их точность может быть недостаточной, и в большинстве случаев, в методах машинного обучения анализ данных происходит после того, как событие уже произошло. Глубокие нейронные сети, являются подразделом машинного обучения. Они способны обрабатывать большие объемы данных и улавливать сложные закономерности, что делает их незаменимыми при решении задач, где традиционные методы неэффективны.

Выбор метода прогнозирования зависит от конкретной задачи и доступных ресурсов, однако для проактивного мониторинга на АОЭ наиболее перспективными являются глубокие нейронные сети, но не стоит отбрасывать статистические методы и методы машинного обучения, так как они могут оказаться достаточно эффективными для решения простых и менее критичных задач.

3) Определить пороговые значения отклонений от нормального состояния системы. Определение пороговых значений требует глубокого анализа данных и экспертного мнения, но это позволит различать нормальные колебания и потенциальные неисправности в работе системы. Важно отметить, что для успешной реализации этого подхода необходима полная обширная информация о работе системы о ее параметрах, условиях эксплуатации и т. д. Также необходимо учитывать, что нормальное состояние системы может меняться со временем, поэтому необходимо периодически переобучать модель, чтобы она оставалась актуальной.

Внедрение всех предложенных этапов позволит повысить точность диагностики и тем самым повысить надежность системы.

Структура системы мониторинга данных.

Основной функцией системы мониторинга является раннее обнаружение тенденций и закономерностей в телеметрии данных. Общая структура системы изображена на рисунке 1.



Рисунок 1. Структура системы мониторинга данных

Рассматриваемая система представляет собой сложную архитектуру, состоящую из нескольких взаимосвязанных модулей, используемых для обработки и анализа массивов данных. В самом начале системы находится модуль агрегации данных, играющий роль главного узла сбора информации. В него поступает телеметрия из АОЭ, которая представляет собой набор данных о состоянии и работе оборудования, процессов и других параметров. Данные поступают с различной частотой и в разных форматах. Одновременно с поступлением данных из потока телеметрии, модуль агрегации занимается сбором информации из лог-файлов, которые представляют собой записи о событиях, произошедших в системе. Эта информация, как правило, уже структурирована, но может находиться в различных файлах и требует объединения в единое целое. Важно отметить, что модуль агрегации не только собирает и структурирует данные, он также предварительно их обрабатывает. В частности, осуществляется временное хранение, которое предоставляет доступ к данным по времени, что является важным инструментом для анализа. Большое количество телеметрии, поступающей в систему, делает необходимым процесс суммирования. Этот процесс позволяет выявить скрытые закономерности в данных, что является важным для интеллектуального анализа. Суммирование данных осуществляется по разным критериям: по времени, по типу события, по типу оборудования. После завершения агрегации и суммирования информация передается в модуль предварительной обработки данных.

Модуль предварительной обработки данных играет ключевую роль в подготовке данных к дальнейшему анализу. Очистка данных является первым шагом в предварительной обработке. Она позволяет удалить из данных ошибочные значения и дубликаты, которые могут исказить результаты анализа. На втором шаге осуществляется выбор данных по заданным критериям. Например, можно выбрать только данные, относящиеся к определенному типу оборудования, или к определенному периоду времени. Третьим шагом является процесс нормализации, который предназначен для приведения данных к единому масштабу. Это позволяет сравнивать данные, полученные из разных источников по разным показателям, независимо от их исходных единиц измерения. На четвертом шаге осуществляется преобразование данных. Это включает в себя различные операции, такие как замена значений, вычисление новых показателей и группировка данных. Извлечение признаков из данных является пятым шагом, это один из самых важных этапов в предварительной обработке. На этом шаге происходит извлечение из данных скрытых закономерностей и формирование новых признаков, которые могут быть использованы для дальнейшего обучения моделей нейронной сети.

На основе полученных данных система генерирует прогнозы, описывающие потенциальные неисправности и их последствия. Полученные результаты прогнозирования передаются в модуль управления для последующего анализа и принятия решений.

Интеллектуальный метод обнаружения аномалий.

Проактивный мониторинг, в котором основной упор сделан на выявление потенциальных проблем, стоит перед выбором между точностью прогнозов и их своевременностью. Точность прогнозов является основным требованием при мониторинге, так как неточные предсказания отнимают время и ресурсы, в то время как реальные проблемы остаются незамеченными. Своевременность прогнозов также играет важную роль в мониторинге, поскольку ранний прогноз дает больше времени на предотвращение проблем. В свою очередь позднее предсказание, более точно описывает состояние системы и предоставляет более детальную информацию. Это связано с тем, что в процессе мониторинга накапливается больше данных, которые позволяют сделать прогноз более точным. Между этими требованиями существует взаимосвязь, поэтому выбор оптимального баланса между точностью и своевременностью зависит от конкретной системы мониторинга и ее целей. Для систем с высокой критичностью приоритетом является своевременность, так как для минимизации рисков необходимо раньше обнаружить потенциальные неисправности, даже если это является неточностью. Системы с меньшей критичностью, могут позволить себе более высокую точность и возможность откладывать реакцию на потенциальную неисправность. Достижение баланса между точностью и своевременностью прогноза является сложной задачей, требующей тщательного анализа конкретных условий работы того или иного объекта.

Для решения поставленной задачи был выбран ансамбль глубоких нейронных сетей, в основе которого лежит модель обучения LSTM (Long Short-Term Memory) [4, 5]. LSTM-модель, является разновидностью рекуррентных нейронных сетей и подходит для обработки временных рядов, где необходимо учитывать зависимости между предыдущими значениями. Ансамбль LSTM-моделей улучшает прогнозируемые способности за счет объединения результатов нескольких моделей, каждая из которых обучена на слегка отличающихся данных. Этот метод позволяет получить более точные оценки правильности прогноза, так как в основе ансамбля лежит принцип коллективного разума. Каждая LSTM модель в ансамбле отдает голос за свой прогноз, и результат получается путем усреднения. При таком подходе, ансамбль как бы сглаживает индивидуальные ошибки каждой модели, что приводит к более надежным результатам. Дополнительным преимуществом ансамбля является возможность анализа

индивидуальных ошибок каждой модели. Анализируя каждую LSTM-модель, можно выявить закономерности и определить, какие типы данных вызывают наибольшие трудности у каждой из моделей. Это позволяет в дальнейшем улучшить процесс обучения, например, добавив в обучающий набор дополнительные данные, которые наиболее часто ошибочно предсказываются. Для обучения ансамбля был применен метод бутстрапа, который заключается в генерации новых обучающих наборов данных путем равномерной выборки с заменой исходного набора данных. Это позволяет создавать множество различных обучающих наборов, на которых каждая LSTM-модель в ансамбле обучается независимо. Такой подход повышает разнообразие обучающих данных и снижает риск переобучения. Алгоритм имеет следующий вид:

$$D = \{(x_1, y_1), (x_2, y_2), (x_3, y_3, \dots, (x_n, y_n))\} - \text{входной набор данных};$$

$$t = 1, \dots, T, \text{ где } T - \text{количество базовых учащихся};$$

$$h_i = \psi(D, D_{bs}) \% D_{bs}, \text{ где } \psi - \text{базовый алгоритм обучения.}$$

Результат:

$$H(x) = \arg \max_{y \in Y} \sum_{t=1}^T I(h_t(x) = y).$$

Подход на основе ансамбля LSTM для обнаружения потенциальных неисправностей.

Предлагаемый метод заключается в том, что ансамбль моделей LSTM на основе данных создает прогноз в каждой потенциальной контрольной точке и обеспечивает оценку надежности этого прогноза. Прогнозы могут быть сделаны на различных этапах мониторинга (часы, дни и месяцы). В этом случае точкой выполнения прогноза является контрольная точка. Подход для обнаружения потенциальных неисправностей на основе ансамбля LSTM изображен на рисунке 2.

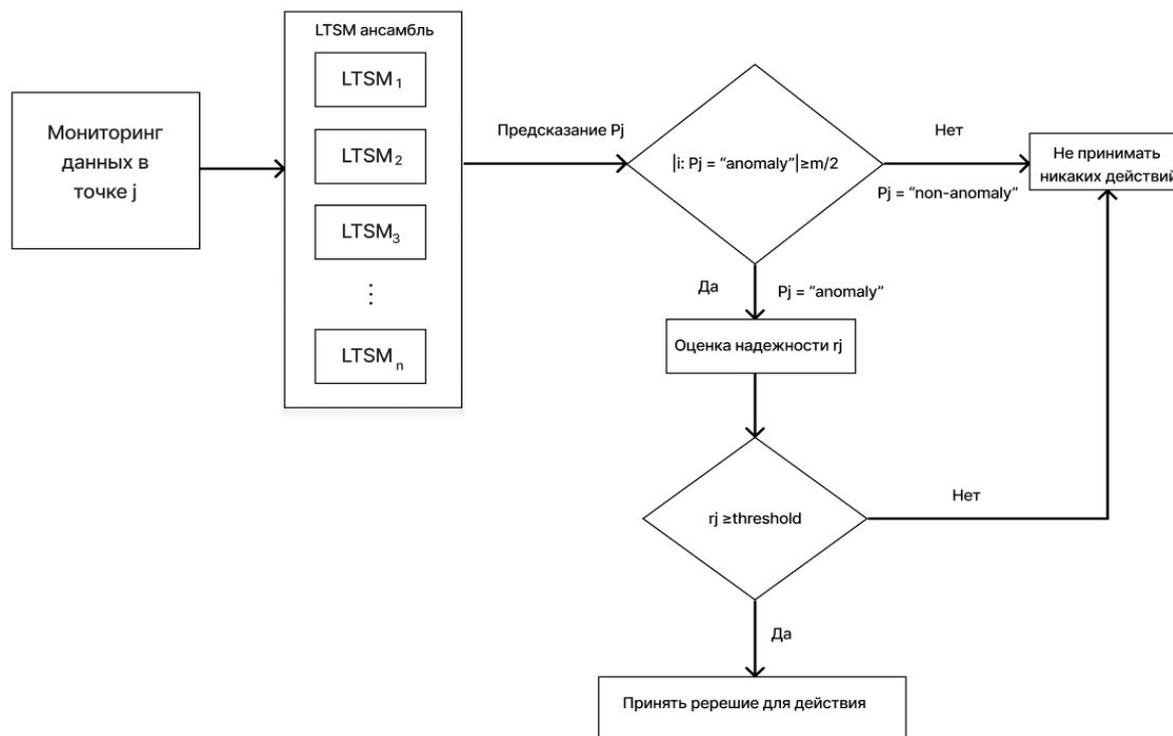


Рисунок 2. Подход для обнаружения потенциальных неисправностей на основе ансамбля LSTM

Если ансамбль LSTM-моделей обнаруживает отклонение от нормального состояния системы, он сигнализирует о потенциальной неисправности. В этом случае предпринимаются меры по ее предотвращению, но перед тем, как предпринимать какие-либо действия, необходимо оценить достоверность прогноза. Для этого устанавливается порог достоверности и если прогноз превышает этот порог, то меры по предотвращению потенциальной неисправности предпринимаются. В случае, если прогноз не превышает порог достоверности, никаких мер не предпринимается. Прогноз для контрольной точки рассчитывается по следующей формуле:

$$P_j = \begin{cases} "anomaly", & |i : P_{ij} = "anomaly"| \geq \frac{m}{2} \\ "non - anomaly", & otherwise \end{cases} \quad (1)$$

Оценка надежности рассчитывается по следующей формуле:

$$r_j = \max_{i=1, \dots, n} \left(\frac{|i : P_{ij} = "anomaly"|}{n}, \frac{|i : P_{ij} = "non-anomaly"|}{n} \right). \quad (2)$$

Для оценки предлагаемого подхода использовался набор потенциальных неисправностей, возникающих на АОЭ. Данные включает в себя:

1. Аварии, связанные с выходом электрических параметров за пределы допустимых значений.
2. Кратковременные замыкания.
3. Выход из строя оборудования.
4. Ложные срабатывания защит.

Модель обучения LSTM реализована на Python 3.8.2 с использованием различных библиотек, включая TensorFlow и Keras. Она состоит из трех слоев LSTM и слоя активации. Использовалась двоичная функция кросс-энтропийных потерь, а также оптимизатор RMSProp, поскольку он дает наиболее стабильные результаты. Для повышения эффективности обучения использовались пакеты по 50 экземпляров, а градиенты вычислялись после каждого пакета. Размерность входных данных была ограничена длиной окна 100, чтобы избежать переобучения модели, использовался слой отсева для упорядочивания активаций. Для генерации выходных данных архитектуры, использовалась сигмоидальная функция активации, выходным значением которой является значение от 0 до 1. Что касается обучения предлагаемой модели, набор был разделен на наборы данных обучения / проверки с соотношением 80% / 20%. Ожидаемая доля потенциальных неисправностей во всем наборе установлена равной 10%, так что в качестве аномалий выбраны 10% наиболее достоверных данных из прогнозов. Максимальное значение среднего среднеквадратичного отклонения (МСЭ) в обучающем наборе принималось в качестве погрешности при прогнозировании состояния в контрольной точке [6].

Выводы. Предлагаемый ансамбль глубоких нейронных сетей на основе моделей обучения LSTM может быть применен в прогнозировании потенциальных неисправностей на АОЭ. Система успешно обнаружила большинство потенциальных неисправностей. Однако, некоторые категории распознавались менее эффективно из-за недостаточного количества данных в обучающей выборке. В частности, к ним относились кратковременные замыкания и ложные срабатывания защит. Модель продемонстрировала хорошую точность в прогнозировании потенциальных неисправностей. Однако постоянный мониторинг данных, проводимый на АОЭ несмотря на свою эффективность, может негативно сказываться на общей производительности системы. Причина кроется в том, что постоянный поток измерений, приводит к значительному увеличению сетевого трафика. Это особенно актуально для

систем с ограниченными ресурсами или высокой загрузкой. В этом случае для оптимизации и повышения эффективности мониторинга возможен переход к мониторингу данных по необходимости. Данный подход предполагает сбор и анализ данных только в тех случаях, когда это действительно необходимо, например, при возникновении отклонений от установленных параметров или в случае проведения плановой диагностики. В настоящее время ведутся исследования по разработке интеллектуальной системы поддержки принятия решений для мониторинга по необходимости. Данная система поможет снизить нагрузку на сетевую инфраструктуру и освободить ресурсы сети для других задач. Разработка подобной системы является сложной задачей, требующей комплексного подхода и глубокого анализа данных. Однако, потенциальные преимущества такой системы значительно превышают сложность ее разработки. В будущем реализация этого подхода позволит улучшить качество мониторинга, сэкономить ресурсы сети и повысить общую эффективность работы АОЭ.

СПИСОК ИСТОЧНИКОВ

1. Ганеева Ж.Г. Определение понятия «мониторинг» в различных сферах его применения / Ж.Г. Ганеева // Вестник Челябинского государственного университета. – 2005. – Т. 8. – № 1. – С. 30-33.
2. Охтилев М.Ю. Теоретические и технологические основы концепции проактивного мониторинга и управления сложными объектами / М.Ю. Охтилев, Б.В. Соколов, Р.М. Юсупов // Известия ЮФУ. Технические науки. – 2015. – № 1 (162). – С. 162-174.
3. Авакьянц А.В. Методы прогнозирования / А.В. Авакьянц, М.Ю. Урубкин // Инновационная наука. – 2017. – № 2-1. – С. 12-14.
4. Староверов Б.А. Реализация глубокого обучения для прогнозирования при помощи ансамбля нейронных сетей / Б.А. Староверов, Р.Н. Хамитов // Известия Тульского государственного университета. Технические науки. – 2023. – Вып. 4. – С. 185-189.
5. Созыкин А.В. Обзор методов обучения глубоких нейронных сетей / А.В. Созыкин // Вестник Южно-Уральского государственного университета. Серия: Вычислительная математика и информатика. – 2017. – Т. 6. – № 3. – С. 28-59.
6. Хайкин С. Нейронные сети: полный курс, 2-е издание / С. Хайкин. – Москва: Издательский дом «Вильямс», 2006. – 1104 с.

ИНФОРМАЦИЯ ОБ АВТОРАХ

Карпухин Дмитрий Николаевич, аспирант кафедры электропривода, автоматике и управления в технических системах, Воронежский государственный технический университет, Воронеж, Россия.

e-mail: karpuxin.1996@mail.ru, *тел.:* +7 (473) 243-77-20

Бурковский Виктор Леонидович, доктор технических наук, профессор, профессор кафедры электропривода, автоматике и управления в технических системах, Воронежский государственный технический университет, Воронеж, Россия.

e-mail: bvl@vorstu.ru, *тел.:* +7 (473) 243-77-20