УДК 004.421

DOI: 10.15827/0236-235X.121.212-217

Дата подачи статьи: 14.06.17

2018. Т. 31. № 1. С. 212–217

ПРОГНОЗИРОВАНИЕ АНОМАЛИЙ В РАБОТЕ НАТЯЖНЫХ УСТРОЙСТВ АГРЕГАТА ПОЛИМЕРНЫХ ПОКРЫТИЙ МЕТАЛЛА В УСЛОВИЯХ МАЛОГО КОЛИЧЕСТВА ОТКАЗОВ

Н.И. Шаханов¹, аспирант, nikshakhanov@gmail.com**И.А. Варфоломеев**¹, к.т.н., доцент, igor.varf@gmail.com**О.В. Юдина**¹, к.т.н., доцент, ovjudina2010@mail.ru**Е.В. Ершов**¹, д.т.н., профессор, eve@chsu.ru

¹ Череповецкий государственный университет,
просп. Луначарского, 5, г. Череповец, 162600, Россия

В статье представлена модель прогнозирования аномалий в работе натяжных устройств агрегата полимерных покрытий металла на основе алгоритма машинного обучения Random Forest при малом количестве прецедентов в обучающей выборке.

Задача алгоритма машинного обучения – спрогнозировать нормальный сигнал для каждого тока электродвигателя, используя значения сигналов в текущий промежуток времени от остальных электродвигателей. Random Forest используется для построения моделей, прогнозирующих нормальную работу оборудования. Обучение модели проводится на данных, считанных с датчиков оборудования в период его эксплуатации, настройка модели осуществляется на данных о предыдущих отказах по натяжным устройствам. Для выявления отказов модель анализирует разность в показаниях фактического и прогнозного сигналов в каждый момент времени с выбранным интервалом.

В статье описаны этапы построения данной модели, представлена схема агрегата полимерных покрытий и дана информация по настройке модели в ПО, специально разработанном для прогнозирования аномалий в работе натяжных устройств агрегата полимерных покрытий металла. Разработан подход, который позволяет быстро и эффективно внедрять созданные модели прогнозирования на предприятии. Для быстрого переобучения модели на языке R и конвертации обученной модели в классы на языке C# (на котором разработано ПО) используется специальный скрипт-транслятор.

Использование разработанной модели позволяет в автоматическом режиме эффективно следить за работой агрегата и информировать диспетчера ремонтного подразделения об обнаруженных аномалиях в работе электродвигателей.

Ключевые слова: агрегат полимерных покрытий металла, прогнозирование аномалий, машинное обучение, случайный лес, дерево решений.

Современное оборудование снабжено необходимым количеством датчиков и приборов контроля, позволяющих отслеживать большое число параметров его работы. При прогнозировании отказов необходимо выбрать наиболее информативные из них. Для решения этой задачи могут быть использованы методы *машинного обучения* (МО). В последнее время прогнозирование отказов оборудования на основе методов МО становится все более актуальным в таких областях, как машиностроительное производство, транспортные системы, промышленность, поскольку они значительно эффективнее методов, основанных на статистическом прогнозировании и т.п. [1, 2]. Однако предложенные к построению прогнозных моделей подходы ориентированы на обучение на прецедентах, что в ряде случаев невозможно применить на практике, например, в ситуациях, когда поломки происходят довольно редко или нет достаточно обширной статистики. Так, на агрегате полимерных покрытий металла можно прогнозировать неисправности в работе приводных электродвигателей роликов у натяжных устройств, используя подход прогнозирования отказов оборудования в условиях малого количества поломок [3].

При данном подходе предлагается применить алгоритм МО Random Forest для построения моделей, прогнозирующих нормальную работу оборудования. Исторические данные используются для обучения модели, а данные о поломках – для коррекции модели. Подход состоит в следующем: осуществляются сбор необходимых данных, обучение модели на данных, соответствующих нормальному режиму работы, прогнозирование работы оборудования и анализ прогнозных и фактических данных с датчиков оборудования для выявления аномалий (потенциальных поломок). Основной принцип выявления отказов оборудования – обнаружение превышения разницы между фактическими и прогнозными значениями за ограниченный период времени (окно). Превышение заданного порога более чем в 95 % случаев свидетельствует об аномалии. Процент превышения порога вычисляется на этапе обучения модели.

Натяжные устройства агрегата полимерных покрытий металла

Агрегат полимерных покрытий металла используется для нанесения лакокрасочного материала на

поверхность оцинкованной полосы. На рисунке 1 схематично показана одна из существующих линий по окраске оцинкованной полосы.

Последовательность операций, осуществляемых на линии окрашивания, следующая. Исходные рулоны подают на разматыватели (см. рис. 1), на гильотинных ножницах обрезают концы полос и участки полосы с дефектами. Для соединения полос используется сшивная машина. Натяжение полосы на участках «ножницы»–«сшивная машина»–«входной петлевой накопитель» производится с помощью натяжных роликов 3 и натяжной станции № 1. Далее полоса попадает в петлевой накопитель, а затем в направляющие ролики и правильную машину, перед и за которой установлены натяжные устройства 8. После правильной машины полоса поступает на участок химической подготовки ее поверхности перед нанесением краски. Она заключается в удалении защитной смазки и механических загрязнений с помощью щелочного раствора моющего средства, хромировании, струйной промывке водой. Сушку полосы проводят горячим воздухом, нагретым до 230 °C в сушильном устройстве.

После прохождения натяжной станции № 2 полоса подвергается окраске в окрасочной камере. Окрашивание полосы происходит в два этапа. Первый слой лакокрасочного материала наносят валковым способом с помощью двух окрасочных головок, а после этого его подвергают сушке при температуре 215–225 °C в печи полимеризации № 1.

Охлаждение полосы производят в ванне 17 деминерализованной водой с воздушным сушильным устройством. Далее полоса поступает в натяжную станцию № 3 и в нижнюю часть окрасочной камеры с тремя окрасочными головками. После этого производят сушку покрытия при температуре 230–240 °C в печи полимеризации № 2, охлаждение полосы в ванне охлаждения с деминерализованной водой и сушку в сушильном устройстве.

В выходной секции полоса проходит выходной накопитель, поступает на гильотинные ножницы, где производятся деление полос на заданную длину, вырезка шва, отбор проб. Далее полосу сматывают в рулон, взвешивают и упаковывают.

В линии применены накопители вертикального типа, в каждом из них полоса образует семь петель [4].

Прогнозирование аномалий на основе алгоритма MO Random Forest

Процедура МО требует достаточного количества данных, соответствующих нормальному режиму работы [5]. Сбор данных, соответствующих тем или иным отказам, требует наблюдения в течение длительного времени эксплуатации, и этих данных может быть критически мало, например, для нового оборудования. Поэтому обучение осуществляется на данных, соответствующих нормальному режиму функционирования объекта исследования (то есть в то время, когда не было поломок и других аномалий).

При реализации такого подхода предлагается использовать алгоритм MO Random Forest для построения моделей, прогнозирующих нормальную работу оборудования. Исторические данные используются для обучения модели, а данные о поломках – для ее коррекции. Подход состоит в следующем: осуществляются сбор необходимых исторических данных, обучение модели на нормальных данных, прогнозирование работы оборудования и анализ прогнозных и фактических данных с датчиков оборудования для выявления аномалий (потенциальных поломок).

Алгоритм MO Random Forest применяется при решении задач прогнозирования в различных отраслях [6]. Он основан на построении большого числа (ансамбля) деревьев решений, каждое из которых строится по выборке, получаемой из исходной обучающей выборки с помощью бутстрепа

(выборки с возвращением). В отличие от классических алгоритмов построения деревьев решений у Random Forest при построении каждого дерева на стадиях расщепления вершин используется только фиксируемое число отбираемых признаков обучающей выборки и строится полное дерево (без усеечения). Классификация осуществляется с помощью голосования классификаторов, определяемых отдельными деревьями, а оценка регрессии – усреднением оценок регрессии всех деревьев. Точность классификации зависит от разнообразия классификаторов ансамбля: чем больше классификаторов, тем выше вероятность корректной классификации.

Плюсы Random Forest: высокая точность классификации, защита от переобучения; легкая параллелизация; обучающая выборка для построения случайного леса может содержать признаки, измеренные в разных шкалах; случайные леса могут использоваться для задач не только классификации и регрессии, но и выявления наиболее информативных признаков, кластеризации, определения прототипов классов и выделения аномальных наблюдений [7, 8].

ПО прогнозированию аномалий натяжных устройств агрегата полимерных покрытий в условиях малого количества отказов

Рассмотрим настройку и применение модели для прогнозирования аномалий в работе натяжных устройств агрегата полимерных покрытий. Настройка модели заключается в определении перечня всех параметров, которые можно собрать с узлов агрегата, в определении выборки с историческими данными для обучения и в обучении модели с последующей настройкой порогов срабатывания. Для натяжных устройств агрегата полимерных покрытий фиксируются данные по токам приводных электродвигателей для каждого из натяжных роликов, эти данные собираются в БД в процессе его эксплуатации (для 13 роликов по всем натяжным устройствам агрегата).

Для формирования входных параметров, обучения, оценки и корректировки модели входные данные разбиваются на три основных периода (рис. 2). Отрезок 1 с выборкой данных из БД используется для обучения модели. Данные на отрезке 2 используются для оценки модели, определяется средняя абсолютная ошибка (MAE – Mean Absolute Error).

Рис. 2. Данные для обучения

Fig. 2. Training Data

Рис. 3. Модель прогнозирования тока j электродвигателя по другим токам электродвигателей: I_i – ток i -го ролика фактический, I_i прогноз – величина тока, прогнозируемая моделью I_i в текущий момент времени

Fig. 3. Model of j electric motor current prediction by other electric motor currents

Для работы модели необходимо задать три параметра, которые определяются на отрезке 3: окно для мониторинга токов, коэффициент для порога срабатывания и допустимый процент превышения порога у всех значений в заданном окне.

Для формирования входных параметров и дальнейшего обучения алгоритма на отрезке 1 отбрасываются все интервалы времени, в которые данные были неполными, зафиксированы отказы в работе роликов или работа любого из роликов осуществлялась в аномальном режиме.

Для каждого ролика строится своя модель. На вход каждой модели подаются токи с остальных роликов, на выходе модель выдает прогнозируемое значение тока по выбранному ролику.

Полученные исторические данные, когда все ролики работали в нормальном режиме, используются для дальнейшего обучения каждой модели. Для всех натяжных устройств агрегата полимерных покрытий будут использоваться 13 моделей для прогнозирования тока для каждого электродвигателя ролика (рис. 3).

Таким образом, для каждой модели для обучения определяется список с данными, содержащими показания по току для каждого из роликов в один и тот же момент времени, в противном случае эти данные для обучения не допускаются. Полученная выборка разделяется на две части – обучающую (75 %) и тестовую (25 %). Обучив модель по алгоритму Random Forest, получаем 13 моделей, которые будут использоваться для прогнозирования.

Каждая из моделей представляет собой 100 бинарных деревьев решений. Пример бинарного дерева решений для подмодели I_1 представлен на рисунке 4.

После построения всех моделей необходимо определить среднюю абсолютную ошибку:

$$MAE = \frac{1}{N} \sum_{t=1}^N |I(t) - \hat{I}(t)|, \quad (1)$$

где N – количество точек из выборки за отрезок 2 (рис. 2); $\hat{I}$ – прогнозное значение в момент времени t ; I – фактическое значение тока в момент времени t .

Таким образом будет определено MAE для каждого из 13 роликов.

Чтобы определить возможность поломки оборудования в период времени, ограниченный окном слежения, необходимо установить порог – процент

превышения разницы прогнозного и фактического сигналов для каждого ролика относительно полученного МАЕ для этого ролика, умноженного на коэффициент (определяющий, во сколько раз допустимо превышение МАЕ). Практика показывает, что превышение порога должно иметь место в течение некоторого времени. При этом порог необходимо установить так, чтобы не допустить ложных срабатываний и минимизировать пропуск реальных поломок. Если имеются данные о поломках, порог устанавливается на основе этих данных и выбирается интервал времени (окно), в течение которого определяется превышение порога.

Таким образом, если будет зафиксировано превышение порога разности прогнозируемого и фактического тока для некоторых точек свыше максимально допустимого процента относительно всех точек окна, осуществляется информирование о том, что оборудование неисправно.

В целом 3-й этап настройки модели заключается в выборе окна мониторинга, максимально допустимого процента точек с превышением порога относительно окна и в подборе коэффициентов превышения МАЕ для каждого ролика, чтобы минимизировать как количество ложных срабатываний модели, так и пропуск зафиксированных поломок из отрезка 3 времени (рис. 2).

На рисунке 5 показан принцип мониторинга, на данном этапе в реальном времени проводится сравнение всех фактических сигналов с датчиков с их прогнозными значениями.

В данном случае на исторических данных показан пример, когда, смещая окно мониторинга по времени, ПО определяет, что большая часть разности прогнозных и фактических значений в этом окне превышает порог, и сигнализирует об аварии. На рисунке видно, что в дальнейшем по данным с БД был зафиксирован отказ.

На рисунке 6 показана реализованная модель в виде ПО для прогнозирования отказов 13 роликов у 6 натяжных устройств агрегата полимерных покрытий № 1.

Данное ПО прогнозирует ток для каждого ролика в автоматическом режиме путем сравнения фактического и прогнозного сигналов, в случае превышения разности в показаниях относительно порога фиксируется ошибка. При обнаружении ошибок за 1 000 секунд более чем в 95 % случаев выявляется отказ в работе агрегата. ПО информирует оператора, если какая-либо модель зафиксирует превышение допустимого порога и время, когда был превышен этот порог. Для обучения и настройки модели первый отрезок времени с данными составлял 2 дня, второй – 1 день и третий – 2 месяца. При тестировании ПО в течение двух месяцев не было допущено ложных срабатываний. Помимо прогнозирования отказов оборудования, данную программу можно использовать для диагностики узлов агрегата [9].

Рис. 6. Выявление аномалии

Fig. 6. Anomaly detection

Результаты

Таким образом, для слежения за работой роликов натяжных устройств агрегата полимерных покрытий было реализовано ПО для прогнозирования их отказов на основе алгоритма МО Random Forest, позволяющее эффективно в автоматическом режиме следить за работой агрегата и информировать диспетчера ремонтного подразделения в случае обнаружения аномалий в работе электродвигателей.

Генерируемая модель в ПО позволяет прогнозировать ток каждого приводного электродвигателя ролика у натяжных устройств. Обучение модели проводится на данных, считанных с датчиков оборудования в период его эксплуатации. Задача алгоритма МО – спрогнозировать нормальный сигнал для каждого тока электродвигателя, используя значения сигналов в текущий промежуток времени от остальных электродвигателей. Для выявления отказов модель анализирует разницу в показаниях фактического и прогнозного сигналов в каждый момент времени с выбранным интервалом. В аналогичных подходах модель выявляет сигнатуры в показаниях датчиков, которые были зафиксированы до и во время предыдущих отказов оборудования. На основе выявления совпадений этих сигнатур с фактическим сигналом и прогнозируется потенциальный отказ.

При подходе прогнозирования отказов оборудования в условиях малого количества поломок данные об отказах используются только для настройки модели, а именно для исключения ложных оповещений об отказе оборудования и минимизации пропусков реальных поломок. Кроме того, разработан подход, который позволяет быстро и эффективно внедрять разработанные модели прогнозирования на предприятии. Для быстрого пере-

обучения модели на языке R [10] и конвертации обученной модели в классы на языке C# (на котором разработано ПО) используется дополнительное ПО «Генератор программного кода моделей Random Forest». ПО, рассмотренное в данной статье, и ПО «Генератор программного кода моделей Random Forest» были зарегистрированы (свид. № 2016616932 от 22.06.2016) в Роспатенте.

Литература

1. Мальцев В. Прогнозная аналитика для эффективного использования оборудования. 2016. URL: http://filearchive.cnews.ru/files/reviews/2016_03_29/2_Maltsev.pdf (дата обращения: 10.05.2017).
2. Oliver M. Predictive maintenance & service (PdMS). Outline and value proposition. 2014. URL: <https://blogs.saphana.com/wp-content/uploads/2014/11/Predictive-Maintenance-Service-Outline-Value-Proposition.pdf> (дата обращения: 10.05.2017).
3. Шаханов Н.И., Варфоломеев И.А., Ершов Е.В., Юдина О.В. Прогнозирование отказов оборудования в условиях малого количества поломок // Вестн. Череповецкого гос. ун-та. 2016. № 6. С. 36–41.
4. Агрегаты для нанесения полимерных покрытий на прокат. URL: <http://dlja-mashinostroitelja.info/2012/11/agregaty-dlja-nanesenija-polimernyh-pokrytij/> (дата обращения: 10.05.2017).
5. Alestra S., Brand C., Burnaev E., Erofeev P., Papanov A., Bordry C., Silveira-Freixo C. Rare event anticipation and degradation trending for aircraft predictive maintenance. Proc. WCCM 2014, ECCM 2014, ECFD 2014, 2014, vol. 11, pp. 6571–6582.
6. Иванова Е.И. Интеллектуальная информационная система для решения задач прогнозирования неисправностей вагонного оборудования на железнодорожном транспорте // Программные продукты и системы. 2015. № 4. С. 231–236.
7. Чистяков С.П. Случайные леса: обзор // Тр. Карельского науч. центра РАН. 2013. № 1. С. 117–136.
8. Пальмов С.В., Денискова А.О. Случайный лес: основные особенности // Наука сегодня: теоретические и практические аспекты. 2017. С. 51–52.
9. Виноградова Л.Н., Данилицкий С.В. Диагностика опор скольжения агрегатов металлургического производства методами неразрушающего контроля // Череповецкие научные чтения. 2013. Ч. 3. С. 39–41.
10. Кабаков Р.И. R в действии. Анализ и визуализация данных на языке R. М.: ДМК Пресс, 2014. С. 33–36.

**ANOMALY PREDICTION IN TENSIONING DEVICES OPERATION OF A COLOR COATING LINE
OF METAL UNDER CONDITIONS OF SMALL NUMBER OF FAILURES**

N.I. Shakhanov¹, *Postgraduate Student, nikshakhanov@gmail.com*

I.A. Varfolomeev¹, *Ph.D. (Engineering), Associate Professor, igor.varf@gmail.com*

O.V. Yudina¹, *Ph.D. (Engineering), Associate Professor, ovyudina2010@mail.ru*

E.V. Ershov¹, *Dr.Sc. (Engineering), Professor, eve@chsu.ru*

¹ *Cherepovets State University, Lunocharskogo Ave. 5, Cherepovets, 162600, Russian Federation*

Abstract. The article presents an anomaly prediction model in tensioning devices operation in a color coating line of metal based on the Random Forest machine-learning algorithm with a minimum of precedents in a training data.

The task of the machine-learning algorithm is to predict a normal signal for each motor current using signal values in the current time interval from other motors. Random Forest is used to build models that predict the normal operation of equipment. The model is trained on the data from equipment sensors during its operation. The model is tuned on the data of previous failures of tensioning devices. In order to detect failures, the model analyzes the difference in the readings of the actual and forecast signals at each time point with the selected interval.

The article describes the stages of constructing the model, presents the scheme of a color coating line, and gives information on tuning the model in software specially designed to predict anomalies in the operation of tensioners of a color coating line for metal. There is the developed approach that allows implementing the developed forecasting models at an enterprise quickly and effectively. There is a special translator script to re-train the model quickly in R language and convert the model to C# classes (in which the software is developed).

The developed model makes it possible to monitor the unit operation effectively and automatically inform a repair unit supervisor in case of anomalies in electric motor operation.

Keywords: color coating line, anomalies prediction, machine learning, random forest, decision tree.

References

1. Maltsev V. *Prognostnaya analitika dlya effektivnogo ispolzovaniya oborudovaniya* [Forecasting Analytics for Efficient Use of Equipment]. IBM Russia Customer Center, 2016. Available at: http://filearchive.cnews.ru/files/reviews/2016_03_29/2_Maltsev.pdf (accessed May 10, 2017).
2. Oliver M. *Predictive Maintenance & Service (PdMS)*. Outline and Value Proposition. 2014. Available at: <https://blogs.saphana.com/wp-content/uploads/2014/11/Predictive-Maintenance-Service-Outline-Value-Proposition.pdf> (accessed May 10, 2017).
3. Shakhanov N.I., Varfolomeev I.A., Ershov E.V., Yudina O.V. Predicting equipment failure in minimal data of breakdowns. *Vestn. Cherepovetskogo gos. un-ta* [Cherepovets State Univ. Bulletin]. 2016, no. 6 (75), pp. 36–41 (in Russ.).
4. *Agregaty dlya naneseniya polimernykh pokryty na prokat* [Aggregates for Polymer Coatings Application for Hire]. Available at: <http://dlja-mashinostroitelja.info/2012/11/agregaty-dlja-nanesenija-polimernyh-pokrytij/> (accessed May 10, 2017).
5. Alestra S., Brand C., Burnaev E., Erofeev P., Papanov A., Bordry C., Silveira-Freixo C. Rare event anticipation and degradation trending for aircraft predictive maintenance. *Proc. 11th World Congress on Computational Mechanics (WCCM 2014), 5th European Conf. on Computational Mechanics (ECCM 2014) and 6th European Conf. on Computational Fluid Dynamics (ECFD 2014)*. 2014, vol. 11, pp. 6571–6582.
6. Ivanova E.I. Intelligent information system for rolling stock carriage equipment failures forecasting on railways. *Programmnye produkty i sistemy* [Software & Systems]. 2015, no. 4, pp. 231–236 (in Russ.).
7. Chistyakov S.P. Random forests: a review. *Tr. Karelskogo nauchnogo tsentra RAN* [Trans. of Karelian Research Centre of RAS]. 2013, no. 2, pp. 117–136 (in Russ.).
8. Palmov S.V., Deniskova A.O. Random forest: main features. *Nauka segodnya: teoreticheskie i prakticheskie aspekty* [Science Today: Theoretical and Practical Aspects]. 2017, pp. 51–52 (in Russ.).
9. Vinogradova L.N., Danilitsky S.V. Diagnostics of sliding bearings of metallurgical production units by NDT methods. *Cherepovetskie nauchnye chteniya-2013* [Cherepovets Scientific Readings-2013]. 2013, part 3, pp. 39–41 (in Russ.).
10. Kabakov R.I. *R v deystvii. Analiz i vizualizatsiya dannykh na yazyke R* [R in Action. Data Analysis and Graphics with R]. Moscow, DMK Press, 2014, pp. 33–36 (in Russ.).