

Прогнозирование оставшегося времени безаварийной работы нефтегазодобывающего оборудования с применением технологии искусственных нейронных сетей

М.Г. Ткаченко

Научно-исследовательский институт многопроцессорных вычислительных систем им А.В.Каляева Южного Федерального Университета, Таганрог

Аннотация: В работе рассмотрен метод, основанный на применении искусственных нейронных сетей (ИНС) в задаче прогнозирования оставшегося времени безаварийной работы (ОВР). Для предварительной обработки телеметрических данных в работе предложено использовать аппроксимацию обобщенной функцией Вейбулла, что позволило уменьшить влияние шумовых факторов. Для повышения точности прогнозирования и определения условий окончания обучения внедрен механизм валидации. Проведено экспериментальное сравнение с известным методом прогнозирования ОВР, свидетельствующее о высокой эффективности предложенных решений в ситуации значительного износа оборудования.

Ключевые слова: искусственная нейронная сеть, прогнозирование, наработка на отказ, время безаварийной работы, распределение Вейбулла, нефтегазодобывающее оборудование.

Введение

Обслуживание оборудования на основе анализа оперативного состояния (АОС) позволяет добиться надежного функционирования в ряде сфер повышенной ответственности, таких как авиакосмическая отрасль, энергетика, добыча полезных ископаемых [1-5]. В системах АОС собираются и обрабатываются электротехнические, физико-химические, акустические и другие параметры функционирования технологических объектов для оценки состояния оборудования и прогнозирования ОВР. На основе полученных значений ОВР осуществляется планирование оптимальных операций по оперативному обслуживанию и ремонту, направленных на сокращение времени простоев и минимизацию стоимости ремонта. Точность определения ОВР является критичным параметром эффективной работы систем АОС.

Существующие методы прогнозирования ОВР можно разделить на две группы: методы, основанные на физической модели, и методы, основанные

на анализе данных телеметрии. Физические методы используют модели износа механических элементов; при успешном построении модели, системы данного класса могут формировать оперативные оценки ОВР с высокой точностью. Однако, построение адекватной физической модели – сложный процесс, и его реализация зачастую невозможна в рамках реального производственного процесса.

Методы, основанные на анализе данных, используют ретроспективные данные телеметрических замеров для формирования оценок ОВР. Известен ряд статистических подходов, среди которых можно выделить работу [1]. Также известно применение скрытых моделей Маркова [6] и других методов [7,8] прогнозирования состояния объектов. Перспективным подходом является применение ИНС, получающих оценки с высокой точностью благодаря свойствам адаптации и нелинейности [2,3,9]. Результатом работы перечисленных методов прогнозирования является значение ряда параметров, используемых для последующей оценки состояния технологического объекта. Однако при практическом применении сложно определить границу возникновения аварийной ситуации на основе анализа получаемых оценок. В работе [10] предложен альтернативный подход с применением ИНС, формирующей прогноз ОВР. В данной статье предложены подходы к повышению точности этого метода.

Модифицированный метод оценки ОВР с применением ИНС

В методе [10] для оценки ОВР используется только один параметр, обрабатываемый ИНС с тремя входами и одним выходом. На входы подаются значения наработки, замеров в текущий и предшествующий моменты времени. Выходом сети является значение ОВР. Однако эта модель не может обрабатывать широкий ряд практических задач, в которых данные мониторинга собираются в произвольные моменты времени, например, в задачах анализа состояния погружного оборудования [2,3]. Более того, в

реальных задачах фиксируется значительное число параметров, характеризующих степень износа оборудования, которые должны быть проанализированы ИНС для получения более точной оценки ОВР.

На основе этих замечаний предлагается модифицировать метод [10]. Предлагаемая модель ИНС содержит входной слой, два скрытых слоя и выходной слой. Необходимость использования двух скрытых слоев определена в результате проведения экспериментов, свидетельствующих о повышении точности оценок при внедрении дополнительного скрытого слоя.

Оригинальный метод, предложенный в [10], характеризуется рядом особенностей: при обучении и построении оценок ИНС используются необработанные данные мониторинга состояния объекта; в процессе обучения не применяется валидация.

Особенности использование необработанных данных

В практических приложениях измерения, получаемые в заданные промежутки времени могут содержать шумовые составляющие, существенно изменяющие значения измеряемых параметров [2,3]. Передача необработанных зашумленных данных в модель ИНС может существенно влиять на точность получаемых оценок ОВР.

Для уменьшения влияния шумов предлагается преобразовывать исходные данные с применением обобщенной функции распределения Вейбулла. Распределение Вейбулла позволяет гибко задавать параметры для представления распределении с различными масштабами и видами.

Предлагается следующее обобщение функции распределения Вейбулла для серий измеряемых параметров:

$$z(t) = Y + K \frac{k}{\delta} \left(\frac{t}{\delta}\right)^{k-1} e^{-\left(\frac{t}{\delta}\right)^k} \quad (1)$$

где t – время работы оборудования, $z(t)$ -полученное аппроксимирующее значение; $\frac{k}{\delta}(\frac{t}{\delta})^{k-1}e^{-(\frac{t}{\delta})^k}$ – функция распределения Вейбулла. Параметр K используется для масштабирования полученных значений, параметр Y используется для задания смещения.

В формуле (1) присутствуют 4 параметра, которые должны быть определены так, чтобы минимизировать среднеквадратичное отклонение (СКО) получаемых оценок от исходных замеров телеметрических параметров. В данной работе для определения оптимальных параметров функции (1) применен генетический алгоритм, дающий хорошие результаты в задаче поиска глобального оптимума [2,4]. Тестирование показало, что применение генетического алгоритма для оптимизации параметров функции (1) на всех наборах данных дало высокую точность получаемой аппроксимации исходной последовательности замеров.

Построение модели ИНС

Предлагаемая модель ИНС использует преобразованные данные телеметрических замеров в качестве входных параметров. В работе используются два последовательных замера, разделенных фиксированным временным интервалом. Хотя такое решение существенно снижает точность работы, применение данных трех и более замеров затруднено из-за большого числа анализируемых параметров, влияющих на размеры ИНС и объем требуемой обучающей выборки.

Цель обучения ИНС состоит в том, чтобы предсказать ОВР на основе имеющихся данных времени работы и телеметрических замеров оперативного состояния. Отношение между данными телеметрии, временем работы и ОВР является сложным и нелинейным. В связи с возможностью моделирования нелинейных связей, аппарат ИНС является мощным инструментом для использования для решения поставленной задачи.

Обучение ИНС

На первом этапе обучения производится поиск параметров обобщенных функций Вейбулла, сглаживающих последовательности значений каждого параметра телеметрических замеров. Поиск параметров производится для каждого объекта в обучающей выборке независимо. Полученные функции аппроксимации используются для обучения ИНС. Критическим моментом при обучении является недопущение переобучения сети. При переобучении факторы зашумленности моделируются самой сетью, что влияет на возможности обобщения и точность получаемых прогнозов. В работе [10] не определены алгоритмы определения момента остановки процесса обучения, обеспечивающие наилучшие результаты работы обученной ИНС. Распространенным подходом к решению этой проблемы является применение валидации в процессе обучения [11]. При этом на каждой итерации обучения ИНС производится вычисление значения СКО для валидационного подмножества обучающей выборки. В процессе обучения значение СКО должно уменьшаться. После прохождения определенного числа итераций значение СКО на валидационном множестве начнет возрастать, что будет соответствовать моменту переобучения ИНС. На этом этапе следует остановить процесс обучения, получив ИНС с хорошими возможностями моделирования и обобщения.

При использовании валидации необходимо сформировать для обучения валидационное множество. В работе [11] рекомендовано использовать в качестве валидационного множества часть (40%) обучающей выборки. При наличии больших объемов обучающих выборок, эта рекомендация позволяет добиться хороших результатов, однако, в работе [2] отмечена низкая точность работы ИНС в ситуации возникновения неисправностей, по которым накоплены малые объемы обучающих данных. Уменьшение объема обучающей выборки за счет выделения подмножества валидации приведет к

существенному снижению качества обучения ИНС. В работе предложено в качестве валидационного множества использовать подмножество не сглаженных замеров в базе телеметрических выборок. Эксперименты показали, что при таком подходе сохраняется точность модели ИНС, обученной на полном объеме имеющихся замеров, а внедрение валидации позволило повысить точность моделирования и робастность получаемых прогнозов.

На основе построенных множеств обучения и валидации, производится обучение ИНС. Для обучения использован алгоритм Левенберга-Маркварда (ЛМ), подробно описанный в работе [11]. Как и ряд других, алгоритм ЛМ не гарантирует получение глобального оптимума при обучении ИНС. Для улучшения получаемых результатов алгоритм запускался несколько раз и лучшая модель ИНС выбирается в соответствии с минимальным значением СКО на валидационном множестве.

Моделирование ИНС на данных телеметрических замеров нефтегазодобывающего оборудования

Рассмотрим результаты анализа базы данных телеметрических замеров работы нефтегазодобывающего оборудования, предоставленные ОАО «Сургутнефтегаз».

В предложенной базе присутствует информация, снимаемая с 280 различных датчиков, размещенных на установках, оборудованных электроцентробежными насосами. Общее число замеров в базе равно 2654420, что соответствует данным, собираемым со 120 объектов в течение одного календарного года. Каждый срез содержит показания всех датчиков, установленных на станциях управления.

Собранные данные использованы для обучения рассмотренных моделей и сравнения результатов их работы по оценке ОВР. Для проверки получаемых оценок выбрана группа из 25% объектов, не включаемых в

обучающую выборку. Данные этих объектов использованы для сравнения точности оценок ОВР, формируемых предложенным методом и модифицированным методом, описанным в работе [10].

В обоих методах ИНС состоят из входного слоя, содержащего 560 узлов (280 на каждый замер), двух скрытых слоев по 190 узлов и выходного слоя, содержащего один узел. На вход ИНС подаются данные двух последовательных замеров, разделенных временным интервалом. Экспериментально установлено, что наибольшую точность оценок была получена при задании временного интервала в 12 часов.

При сравнении результатов точность определения ОВР рассчитывалась по формуле

$$p = \frac{1}{N} \sum_i \left(\frac{|t_i^{ann} - t_i^r|}{t_i^r} \right), \quad (2)$$

где N – число записей в проверочном множестве; t_i^{ann} – прогноз ОВР, полученный ИНС; t_i^r – ОВР, вычисленное по базе данных телеметрии.

Для каждого объекта проверочного множества последовательно вычислены значения ошибок p в соответствии с формулой (2). Полученные значения разделены на три группы: p_{all} – точность оценок ОВР для всей группы измерений, p_{less5} – точность оценок для случаев, характеризуемых ОВР менее 5 календарных дней; p_{more5} – точность оценок для случаев, характеризуемых ОВР в 5 и более календарных дней.

Для проверки множество тестовых данных было разбито на 5 групп в соответствии с типом зафиксированной неисправности, по каждой группе была рассчитана точность получаемых прогнозов модифицированным методом [10] и методом, предложенным автором. Результаты полученных оценок сведены в таблицу №1.

Таблица № 1

Точность определения ОВР на экспериментальных данных

№ эксперимента	Средние ошибки прогнозов		
	P_{all}	P_{less5}	P_{more5}
1	0.264	0.141	0.268
2	0.281	0.17	0.292
3	0.28	0.152	0.291
4	0.28	0.152	0.291
5	0.372	0.237	0.39
Среднее значение	0.288	0.178	0.297
Среднее значение (метод [10])	0.32	0.28	0.327

Сравнение результатов эксперимента свидетельствует о незначительном увеличении точности получаемых долгосрочных оценок. При этом можно отметить существенное повышение точности оценки ситуаций, характеризуемых ОВР менее 5 дней. Стоит отметить, что точные оценки ОВР в моменты значительной наработки наиболее важны, так как именно они используются для формирования плана работ по оперативному ремонту и обслуживанию оборудования.

Выводы

В данной работе представлен модифицированный метод повышения точности определения ОВР с применением технологии ИНС. Полученная модель ИНС принимает в качестве входных параметров два последовательных телеметрических замера и выдает оценку ОВР исследуемой системы. Для уменьшения эффектов зашумления в работе применена обобщенная функция Вейбулла. Исходные данные используются в качестве опорных точек для подбора параметров обобщенной функции, значения которой используются для обучения ИНС.

Для повышения точности обучения и определения условия окончания обучения предложено использование механизма валидации. Предлагаемая модель ИНС использует в качестве валидационного множество

необработанных данных исходной обучающей выборки телеметрических замеров.

Проведенное экспериментальное исследование свидетельствует о высокой точности получаемых оценок ОВР, что позволяет внедрить предложенный метод в ряд автоматизированных систем нефтегазодобывающих управлений.

Литература

1. Lin D., Banjevic D., & Jardine A. Using principal components in a proportional hazards model with applications in condition-based maintenance. *Journal of the Operational Research Society*, 57, 2006, p.910–919.
2. Ткаченко М.Г. Применение искусственной нейронной сети, оптимизированной генетическим алгоритмом, в задаче анализа состояния технологического оборудования нефтегазодобывающей промышленности// *известия Южного федерального университета. Технические науки*. 2014., №7(156), с. 262-270.
3. Коровин Я.С., Ткаченко М.Г., Кононов С.В. Оперативная диагностика состояния нефтепромыслового оборудования на основе технологий интеллектуальной обработки данных// *Нефтяное хозяйство*, 2012, №9, с.116-119.
4. Liao H., Elsayed, E., & Chan, L. Maintenance of continuously monitored degrading systems. *European Journal of Operational Research*, 175, 2006, p.821–835.
5. Inman D., Farrar C. *Damage prognosis: For aerospace, civil and mechanical systems*. NY: Wiley, 2006, p.122-143.
6. Dong M. Hidden semi-Markov model-based methodology for multi-sensor equipment health diagnosis and prognosis. *European Journal of Operational Research*, 178, 2007, p. 858–878.

7. Клевцов С.И. Моделирование алгоритма краткосрочного прогнозирования изменения быстроизменяющейся физической величины в реальном времени//Инженерный вестник Дона, 2012, №3 URL: ivdon.ru/magazine/archive/n3y2012/920.

8. Р.С. Кузнецов, Ю.В. Тимофеев, Н.А. Смирнов, М.С. Тютяев, А.П. Черкис, Н.Л. Щербакова. Механизмы вычислительного интеллекта при решении задачи автоматизации прогнозирования электроэнергии//Инженерный вестник Дона, 2012, №2 URL: ivdon.ru/ru/magazine/archive/n2y2012/823

9. Tian Z., & Zuo M. Health condition prognostics of gears using a recurrent neural network approach. In Proceedings of the reliability and maintainability symposium, Fort Worth, 2009, p.432-437.

10. Wu S. J., Gebrael N., Lawley M. A., & Yih Y. A neural network integrated decision support system for condition-based optimal predictive maintenance policy. IEEE Transactions on Systems Man and Cybernetics Part A: Systems and Humans, 37, 2007, p. 226–236

11. Rojas, R. Neural networks: A system introduction. Berlin: Springer, 1996, p.509.

References

1. Lin D., Banjevic D., & Jardine A. Using principal components in a proportional hazards model with applications in condition-based maintenance. Journal of the Operational Research Society, 57, 2006, p.910–919.

2. Tkachenko M. Izvestia Uznogo federalnogo universiteta. 2014. №7(156). pp. 262-270.

3. Korovin Y., Tkachenko M. Kononov S. Neftyanoe hozyaistvo. 2012. №9. pp.116-119.

4. Liao H., Elsayed, E., & Chan, L. Maintenance of continuously monitored degrading systems. *European Journal of Operational Research*, 175, 2006, p.821–835.
5. Inman D., Farrar C. *Damage prognosis: For aerospace, civil and mechanical systems*. NY: Wiley, 2006, p.122-143.
6. Dong M. Hidden semi-Markov model-based methodology for multi-sensor equipment health diagnosis and prognosis. *European Journal of Operational Research*, 178, 2007, p. 858–878.
7. Klevtcov.S.I. *Inženernyj vestnik Dona (Rus)*, 2012, №3 URL: ivdon.ru/magazine/archive/n3y2012/920.
8. Kuznecov R.S., Timofeev U.V., Smirnov N.A., Tulaev M.S., Cherkis A.P., Sherbakova N.L. *Inženernyj vestnik Dona (Rus)*, 2012, №2 URL: ivdon.ru/ru/magazine/archive/n2y2012/823
9. Tian Z., & Zuo M. Health condition prognostics of gears using a recurrent neural network approach. In *Proceedings of the reliability and maintainability symposium, Fort Worth, 2009*, p.432-437.
10. Wu S. J., Gebrael N., Lawley M. A., & Yih Y. A neural network integrated decision support system for condition-based optimal predictive maintenance policy. *IEEE Transactions on Systems Man and Cybernetics Part A: Systems and Humans*, 37, 2007, p. 226–236
11. Rojas, R. *Neural networks: A system introduction*. Berlin: Springer, 1996, p.509.