

УДК 621.438, 004.855

ББК 3363.3, 32.813, 22.18

ПОВЫШЕНИЕ РОБАСТНОСТИ НЕЙРОСЕТЕВОЙ МОДЕЛИ МОНИТОРИНГА ГТД НА ОСНОВЕ РЕДУКЦИИ

Светлана Владимировна Новикова

доктор технических наук, профессор

Казанский национальный исследовательский технический университет

им. А. Н. Туполева-КАИ

Казань, Россия

sweta72@bk.ru

Эльмира Шамильевна Кремлева

Казанский национальный исследовательский технический университет

им. А. Н. Туполева-КАИ

Казань, Россия

e-smile29.04@mail.ru

В статье представлен способ повышения робастности нейросетевой модели мониторинга газотурбинного двигателя при стендовых испытаниях за счет уменьшения количества малозначащих связей нейронной сети (нейросетевой редукции). Способ основан на преобразовании задачи обучения нейросети к задаче многокритериальной оптимизации, включающей в себя критерий минимизации ошибки и критерий минимизации абсолютных значений весовых связей нейросети. Последнее требование приводит к выделению малозначащих связей, которые могут быть удалены без потери точности. В результате, способность модели к обобщению значительно возрастает, увеличивается робастность, уменьшается ошибка расчета мониторируемых параметров.

Ключевые слова: робастность модели, нейронные сети, стендовые испытания, мониторинг ГТД, редукция нейронной сети.

UDK 621.438, 004.855

ВБК 3363.3, 32.813, 22.18

REDUCTION –BASED INCREASE OF ROBUSTNESS OF A NEURAL-NET MODEL OF GAS-TURBINE ENGINE MONITORING

Svetlana Vladimirovna Novikova

Doctor of Technical Sciences, professor

Kazan National Research Technical University after A.N.Tupolev-KAI,

Kazan, Russia

sweta72@bk.ru

Jel'mira Shamil'evna Kremleva

Kazan National Research Technical University after A.N.Tupolev-KAI,

Kazan, Russia

e-smile29.04@mail.ru

The article presents a technique of increasing the robustness of a neural-net model of gas-turbine engine monitoring during stand tests due to decreasing the number of insignificant links in the neural network (neural network reduction). The technique is based on converting the problem of neural network learning into that of multi-criteria optimization which includes the error minimization criterion and the criterion of minimizing the absolute values of the weight links of a neural network. The latter requirement leads to reveal of insignificant links which can be deleted without any loss of accuracy. As a result, the ability of the model to summarize increases greatly, robustness increases as well, the calculation error of monitored parameters decreases.

Key words: model robustness, neural networks, stand tests, monitoring of a gas-turbine engine, reduction of a neural network.

Проектирование газотурбинного двигателя (ГТД) включает в себя этап стендовых испытаний, на котором контролируются технологические параметры ГТД в различных режимах, прямо или косвенно. Так как протекающие в двигателе процессы носят нестационарный характер, его параметры зависят от внешних условий, причем вид такой зависимости часто не определен, состав измеряемых параметров в частном случае может быть неполным и т. д., то задачу контроля состояния ГТД можно отнести к задачам в условиях неполной и нечеткой информации [Новикова, 2006, с. 37-38]. Подобные задачи в настоящее время успешно решаются при помощи интеллектуальных нейросетевых систем и моделей [Жернаков, 2001; Емалетдинова, 2013].

Робастность модели, то есть устойчивость результатов моделирования к возмущениям входных данных, является одним из важнейших требований при моделировании технических систем. При нейросетевом моделировании понятие робастности часто заменяется специфическим термином «способность к обобщению», буквально означающем сохранение сетью точности вычислений для наборов данных, не участвующих в непосредственном обучении. Способность к обобщению достигается правильной настройкой синаптических весов в процессе обучения.

Оценку робастности или способности к обобщению нейросетевой модели будем проводить на основе вычисления специфической меры сложности сети, обозначаемой как $VCdim$, и получившей название «мера Вапника-Червоненкиса». В работе [Vapnik, 1971, P. 270] доказано, чем меньше значение меры, тем выше робастность сети. Доказано также, что значение $VCdim$ обратно пропорционально количеству связывающих нейроны синаптических весов. Следовательно, для увеличения робастности модели требуется уменьшить количество межнейронных связей при сохранении точности вычислений.

Для построения нейросетевой модели необходимо обладать набором измеренных параметров моделируемого объекта в виде кортежей «вход-выход». Используется следующая информация, полученная в процессе стендовых испытаний ГТД [Novikova, 2016, pp. 263-265]:

- $n_1(t)$ – частота вращения ротора низкого давления в момент t ;
- $n_1(t-i*\Delta t)$ – частота вращения ротора низкого давления с временной задержкой $i*\Delta t$, $i=1..L/\Delta t$ (L – ширина временного окна);
- $n_2(t)$ – частота вращения ротора высокого давления в момент t ;
- $n_2(t-i*\Delta t)$ – частота вращения ротора высокого давления с временной задержкой $i*\Delta t$, $i=1..L/\Delta t$ (L – ширина временного окна);
- $T_4^*(t)$ – температура газов за турбиной в момент t ;
- $T_4^*(t-i*\Delta t)$ – частота вращения ротора высокого давления с временной задержкой $i*\Delta t$, $i=1..L/\Delta t$ (L – ширина временного окна).

Описанные параметры используются в качестве входной информации. Выходной информацией является режим работы ГТД, условно обозначенный как:

- 0 – установившийся режим;
- 1 – переходной режим;
- 2 – режим разгона;
- 3 – режим дросселирования.

Для выявления оптимальной структуры и топологии нейронной сети было проведено несколько разведывательных экспериментов. В результате был сделан выбор в пользу многослойного персептрона с одним скрытым слоем. Входной слой содержит шесть нейронов по количеству входных параметров, выходной – единственный нейрон (вид режима функционирования ГТД). Функции активации слоев сети выбирались из условия их непрерывности и дифференцируемости, так как именно такой выбор функций гарантирует, что построенная нейронная сеть будет являться универсальным аппроксиматором [Осовский, 2002, с. 50-55], и может моделировать зависимости любой

сложности. В результате, функция активации скрытого слоя определена в виде гиперболического тангенса, выходного слоя – линейной функции.

Тогда функционирование обученной нейросетевой модели с выбранными характеристиками описывает следующая формула:

$$y = \left(\sum_{j=1}^M w_{j1}^{(2)} \left(\frac{e^{\frac{2 \sum_{i=1}^N w_{ij}^{(1)} x_j - 1}{2 \sum_{i=1}^N w_{ij}^{(1)} x_j + 1}} - 1}{e^{\frac{2 \sum_{i=1}^N w_{ij}^{(1)} x_j - 1}{2 \sum_{i=1}^N w_{ij}^{(1)} x_j + 1}} + 1} \right) - 1 \right) \quad (1)$$

Здесь:

N – количество нейронов входного слоя;

M – количество нейронов скрытого слоя;

x_j – входные параметры модели;

y – выходной параметр модели;

$w_{ij}^{(1)}$ – синаптические веса скрытого (первого) слоя $i = \overline{1, N}$, $j = \overline{1, M}$, определенные на этапе обучения;

$w_{j1}^{(2)}$ – синаптические веса выходного (второго) слоя, определенные на этапе обучения;

Редукцию нейросетевой модели будем проводить на этапе обучения, включив в качестве критерия успешности обучения дополнительное условие: значения весов всех слоев сети в процессе обучения должны быть скорректированы так, чтобы:

- 1) функция ошибки сети на обучающем множестве была минимальна;
- 2) абсолютные значения синаптических коэффициентов были минимальны (в дальнейшем это приведет к их исключению и упрощению структуры нейросети).

Тогда процесс обучения описывает двухкритериальной оптимизационной задачей вида:

$$E^{(0)(l)}(w) = \frac{1}{2} \left(\left(\sum_{j=1}^M w_{j1}^{(2)} \left(\frac{e^{\frac{2 \sum_{i=1}^N w_{ij}^{(1)} x_j^{(l)} - 1}{2 \sum_{i=1}^N w_{ij}^{(1)} x_j^{(l)} + 1}} - 1}{e^{\frac{2 \sum_{i=1}^N w_{ij}^{(1)} x_j^{(l)} - 1}{2 \sum_{i=1}^N w_{ij}^{(1)} x_j^{(l)} + 1}} + 1} \right) - 1 \right) - d^{(l)} \right)^2 \Rightarrow \min \quad (2)$$

$$E^{(1)}(w) = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^M \left(\frac{(w_{ij}^{(1)})^2}{1 + \sum_{k=1}^M (w_{ik}^{(1)})^2} \right) + \frac{1}{2} \sum_{j=1}^M \left(\frac{(w_{j1}^{(2)})^2}{1 + \sum_{k=1}^M (w_{kj1}^{(2)})^2} \right) \quad (3)$$

здесь $l = \overline{1, q}$ - номер обучающего примера, $d^{(l)}$ – эталонный выходной вектор с номером l . В представленной модели условие (2) отвечает за минимизацию ошибки на обучающем множестве, а (3) – за уменьшение абсолютных значений весов.

Многокритериальная задача (2)-(3) на каждой итерации обучения решается методом последовательных уступок, при этом условие (1) считается важнее, чем условие (2).

Для решения задачи редукции нейросетевой модели был разработан алгоритм на основе многокритериального обучения:

1. По имеющемуся массиву исходных данных задать:
 - 1.1. q – количество обучающих примеров;
 - 1.2. σ – порог отсечения малозначащих весов;
 - 1.3. ε – величина ошибки нейросетевой модели;
 - 1.4. MAX – предельное количество эпох обучения.
2. Счетчик эпох обучения установить в 1: $k=1$.
3. Положить номер обучающего примера l равным 1.
4. Для l -того обучающего примера сформулировать многокритериальную оптимизационную задачу вида (2)-(3).
5. Решить оптимизационную задачу (2). В качестве численного алгоритма минимизации выбрать метод переменной метрики с моментом [Nocedal, 2006].
 - 5.1. Найти вектор p_n направления оптимизации для синаптических весов выходного слоя $w_{j1}^{(2)n}$, $j = \overline{1, M}$ методом переменной метрики.
 - 5.2. Определить шаг η_n для коррекции весов выходного слоя в направлении p_n методом направленной минимизации.
 - 5.3. Пересчитать значения синаптических весов выходного слоя:

$$\tilde{w}_{j1}^{(2)n+1} = w_{j1}^{(2)n} + \eta_n p_n + \alpha(w_{j1}^{(2)n} - w_{j1}^{(2)n-1})$$

где $\alpha \in [0,1]$ – коэффициент момента.

5.4. Найти вектор p_n направления минимизации для весов скрытого слоя

$w_{ij}^{(1)n}$ $i = \overline{1, N}$, $j = \overline{1, M}$ методом переменной метрики.

5.5. Определить шаг η_n для коррекции весов скрытого слоя в направлении p_n методом направленной минимизации.

5.6. Пересчитать значения весов скрытого слоя по формуле:

$$\tilde{w}_{ij}^{(1)n+1} = w_{ij}^{(1)n} + \eta_n p_n + \alpha(w_{ij}^{(1)n} - w_{ij}^{(1)n-1})$$

6. Решить оптимизационную задачу (3). В качестве численного алгоритма минимизации выбрать метод градиентного спуска [Мурга, 2013, с. 96-98]. В качестве начальной точки работы алгоритма принять оптимальную точку, найденную на шаге 5 в процессе минимизации критерия $E^{(0)(l)}(w)$.

6.1. Пересчет весов выходного слоя:

$$w_{j1}^{(2)n+1} = \tilde{w}_{j1}^{(2)n+1} \left[1 - \eta \frac{1 + 2 \sum_{k \neq j}^M (\tilde{w}_{k1}^{(2)n+1})^2}{\left[1 + \sum_{k=1}^M (\tilde{w}_{k1}^{(2)n+1})^2 \right]^2} \right], \quad j = \overline{1, M}$$

6.2. Пересчет весов скрытого слоя:

$$w_{ij}^{(1)n+1} = \tilde{w}_{ij}^{(1)n+1} \left[1 - \eta \frac{1 + 2 \sum_{k \neq j}^{25} (\tilde{w}_{ik}^{(1)n+1})^2}{\left[1 + \sum_k (\tilde{w}_{ik}^{(1)n+1})^2 \right]^2} \right], \quad i = \overline{1, N}, \quad j = \overline{1, M}$$

7. Если $l < q$, то $l = l + 1$. Возврат к шагу 4. Иначе – к шагу 8.

8. Редуцировать нейросетевую модель:

8.1. Синаптические связи с весом меньше σ по абсолютному значению считать незначимыми и удалить из сети;

8.2. Если в результате удаления синаптических связей нейрон скрытого слоя окажется несвязанным с входными или выходным нейронами, удалить несвязанный нейрон.

9. Завершить вычисления, если:

а) Ошибка сети $E^{(0)}(w)$ меньше заданной погрешности ε :

$$E^{(0)}(w) = \frac{1}{2} \sum_{l=1}^q (y^{(l)}(w) - d^{(l)})^2 \leq \varepsilon \quad (4)$$

б) Достигнуто максимальное количество эпох обучения: $k=\text{MAX}$.

Иначе вернуться к шагу 3.

Для сравнения эффективности разработанного алгоритма с существующими методиками редукции нейронных сетей были выбраны следующие:

- метод OBD (Optimal Brain Damage);
- метод штрафной функции.

Далее были проведены вычислительные эксперименты, в ходе которых нейронные сети изначально одинаковой топологии подвергались редукции тремя методами (предложенным алгоритмом и двумя стандартными). В результате эксперименты продемонстрировали преимущество предложенного алгоритма по точности как на обучающем, так и на тестовом множестве (Табл. 1).

Таблица 1 – Результаты работы алгоритма в сравнении со стандартными методами

Метод редукции	Ошибка обучения (средняя)	Ошибка обобщения (средняя)	Количество эпох обучения
Разработанный алгоритм	1%	1,5%	15 198
ODB	4%	5%	20 000
Метод штрафной функции	2%	3%	20 000

Такой результат говорит о том, что алгоритм повышает робастность модели без потери точности и, как следствие, уменьшает ошибку при расчете контролируемых параметров ГТД. В результате редукции будет снижено

значение коэффициента меры сложности сети $VCdim$, и тем самым повышена робастность нейросетевой модели ГТД.

Выводы: разработанный алгоритм редукции нейронных сетей повышает робастность модели, построенной на ее основе. Алгоритм превосходит стандартные методы по точности и скорости обучения.

Библиографический список

1. *Емалетдинова Л. Ю.* Алгоритмы построения нейросетевой модели динамического объекта регулирования и настройки параметров пид-регулятора / Л. Ю. Емалетдинова, Е. Д. Царегородцева // Известия высших учебных заведений. Авиационная техника. 2013. № 3. С. 27–33.
2. *Жернаков С. В.* Применение нейросетей для отладки параметров ГТД в процессе эксплуатации // Нейрокомпьютеры и их применение: VII Всеросс. конф. с междунар. участием. М.: ИПУ, 2001. С. 161–165.
3. *Мурга О. К.* Методы оптимизации: учеб. пособие. – 2-е изд., испр. и доп. / О. К. Мурга, А. А. Еремеева. Казань: Изд-во КНИТУ-КАИ, 2013. 189 с. ISBN 978- 5-7579-1230-1
4. *Новикова С. В.* Применение алгоритма адаптивной коррекции для регулирования газотурбинного двигателя в условиях помех // Известия высших учебных заведений. Авиационная техника. Казань, Изд-во КГТУ им. Туполева, 2006. №4. С. 37–40
5. *Осовский С.* Нейронные сети для обработки информации / Пер. с польского И. Д. Рудинского. М.: Финансы и статистика, 2002. 344 с.
6. *Nocedal George; Wright Stephen J.* Numerical Optimization. 2nd edition. USA: Springer, 2006. ISBN 978-0-387-30303-1.
7. *Novikova S. V.* Structural Optimization of the Neural Network Model for the Gas Turbine Engine Monitoring // Russian Aeronautics (Iz. VUZ), 2016. Vol. 59, No. 2, pp. 263-270.- ISSN 1068-7998. DOI 10.3103/S1068799816020185.
8. *Vapnik V. N., Chervonenkis A.* On the uniform convergence of relative frequencies of events to their probabilities. //Theory of Probability and its Applications, 1971, Vol.16, Pp. 264-280.

References

1. *Emaletdinova L. Ju.* Algorithms for constructing a neuro-net model of a dynamic control object and settings of pid-regulator parameters / L. Ju. Emaletdinova, E. D. Caregorodceva // Proceedings of higher educational establishments. Aviation engineering. 2013. № 3. P. 27–33. (in Russian)
2. *Murga O.K.* Optimization methods: text-book. – 2nd edition / O. K. Murga, A. A. Ereemeeva. Kazan: KNRTU-KAI, 2013. 189 p. ISBN 978- 5-7579-1230-1. (in Russian)
3. *Nocedal George; Wright Stephen J.* Numerical Optimization. 2nd edition. USA: Springer, 2006. ISBN 978-0-387-30303-1. (in English)
4. *Novikova S. V.* Structural Optimization of the Neural Network Model for the Gas Turbine Engine Monitoring // Russian Aeronautics (Iz. VUZ), 2016. Vol. 59, No. 2, pp. 263-270. - ISSN 1068-7998. DOI 10.3103/S1068799816020185. (in English)
5. *Novikova S.V.* Application of the adaptive correction algorithm for adjusting a gas-turbine engine under noise conditions // Proceedings of higher educational establishments. Aviation engineering. Kazan, KSTU named after Tupolev, 2006. №4. P. 37–40. (in Russian)
6. *Osovskij S.* Neural networks for information processing / Translated from Polish by I.D. Rudinskij. M.: Finansy i statistika, 2002. 344 p. (in Russian)
7. *Vapnik V. N., Chervonenkis A.* On the uniform convergence of relative frequencies of events to their probabilities. //Theory of Probability and its Applications, 1971, Vol.16, Pp. 264-280 (in English)
8. *Zhernakov S. V.* Application of neural networks for fine-tuning of GTE parameters during operation // Neurocomputers and their application: 7th All-Russian international conference. M.: IPU, 2001. P. 161–165. (in Russian)