УДК 656.259.1

САЯПИНА И.А., аспирант (УкрГАЖТ)

Нейросетевое моделирование аппаратуры рельсовых цепей

С целью прогнозирования и диагностики состояния оборудования рельсовой цепи разработана нейросетевая модель рельсовой цепи. Рассмотрен метод обучения нейронной сети на основе алгоритма Левенберга-Марквардта. Приведены результаты обучения и пример функционирования модели.

Ключевые слова*: нейронная сеть, рельсовая цепь, нейросетевое моделирование, алгоритм Левенберга-Марквардта, обучение нейронной сети.*

Введение

В настоящее время в связи с повышением скоростей движения поездов актуальным становится вопрос модернизации и совершенствования систем, отвечающих за безопасность процесса перевозок.

На железных дорогах Украины основным путевым датчиком являются рельсовые цепи (РЦ). Большинство отказов РЦ приходится на изолирующие стыки. Решить данную проблему можно внедрением бесстыковых рельсовых цепей. Дальнейшее повышение безопасности систем на основе РЦ возможно за счет совершенствования технического обслуживания и использования автоматизированных систем диагностики и прогнозирования отказов рельсовых цепей. Для этих целей может быть использовано нейросетевое моделирование, имеющее ряд достоинств [1]:

- способность нейросети (НС) самообучаться, а также с помощью обобщений получать обоснованные результаты при поступлении данных, не встречавшихся при обучении;

- использование НС для решения задач, закономерности развития которых неизвестны, достигается благодаря способности обучаться на множестве примеров;

- адаптивность НС проявляется в способности нейронов адаптировать свои синаптические веса к изменениям окружающей среды, т. е. их можно легко переобучить для работы в условиях некоторых колебаний параметров окружающей среды;

- параллельная обработка информации нейронами позволяет достичь высокого быстродействия;

- отказоустойчивость аппаратной реализации НС достигается за счет распределенной структуры НС, т. к. при повреждении отдельных нейронов производительность системы падает незначительно;

- универсальность НС выражается в том, что одно и то же проектное решение НС может быть применено ко многим предметным областям вследствие единообразия анализа и проектирования НС.

Анализ публикаций

В статье [2] рассмотрена возможность применения систем искусственного интеллекта для прогнозирования параметров качества машин. Показано, что задача прогнозирования с использованием нейронных сетей сводится к задаче аппроксимации многомерных функций, то есть к задаче построение многомерного отображения. В материале [3] представлено моделирование сложных нелинейных систем на основе аппарата нейронных сетей. В источнике [4] отмечены преимущества нейросетевого подхода в решении вопросов построения устойчивых приближенных математических моделей сложных систем с распределенными параметрами по разнородной информации с уточняемыми данными.

Основная часть

Ввиду достоинств нейронных сетей поставлена задача создания нейросетевой модели рельсовой цепи. Для ее реализации использована среда программирования Matlab, в частности ее пакет расширения Neural Network Toolbox, содержащий средства для проектирования, моделирования, разработки и визуализации нейронных сетей.

Благодаря свойству адаптивности НС данная модель позволила дополнительно учесть факторы, влияющие на работу РЦ, такие как колебания сопротивления изоляции рельсовой линии ($R_{из}$) и температура окружающей среды (t^o). Как следствие, входной обучающий сигнал состоит из трех составляющих: значения параметра сопротивления изоляции, температуры окружающей среды и, собственно, самого сигнала контроля состояния РЦ (рис. 1).

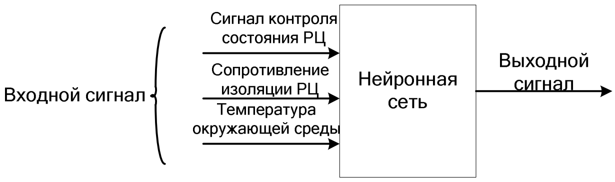

Рис. 1.Структурная схема нейросетевой модели РЦ

Искусственная нейронная сеть представляет собой математическую модель, которую можно рассматривать, как систему взаимодействующих между собой и взаимосвязанных простых процессоров, называемых нейронами. Их создание требует большой вычислительной мощности, поэтому для этого используется компьютеры со специальным программным обеспечением, называемым нейросимулятором или нейропакетом. Задачей данных программ является имитация работы нейросетей.

В качестве структуры была выбрана сеть прямого распространения с одним скрытым слоем, обучаемая с использованием обратного распространения ошибки (рис. 2).

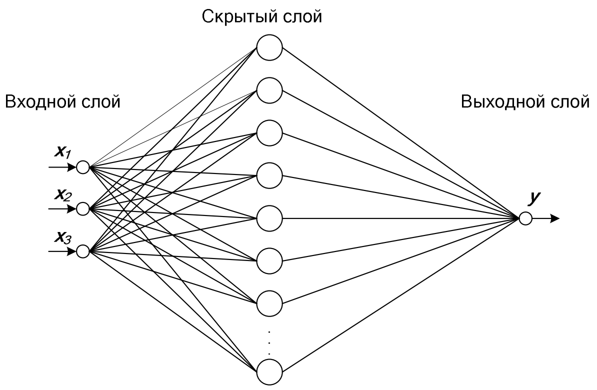

Рис. 2. Сеть прямого распространения

Каждый нейрон скрытого слоя имеет 3 синаптические связи, каждая из которых характеризуется своим синаптическим весом w_i, один аксон, сумматор и активационную функцию.

В данном случае под процессом обучения НС подразумевается нахождение параметров регрессионной модели. [5] Регрессионная модель $f(\boldsymbol{w},\boldsymbol{x})$ – это параметрическое семейство функций, задающее отображение $f:\ \boldsymbol{W}\times\boldsymbol{X}\to\boldsymbol{Y}$,

где $\boldsymbol{w}\in\boldsymbol{W}$ – пространство параметров,

$\boldsymbol{x}\in\ \boldsymbol{X}$ – пространство свободных элементов,

$\boldsymbol{Y}$ – пространство зависимых элементов.

Для обучения нейронной сети используется Байесовская регуляризация на основе алгоритма Левенберга-Марквардта, предназначенного для оптимизации параметров нелинейных регрессионных моделей. [6, 7] Критерием при оптимизации является среднеквадратическая ошибка модели на обучающей выборке. Работа алгоритма направлена на последовательное приближение начальных значений параметров к искомому локальному оптимуму.

В данном случае регрессионной моделью $f(\boldsymbol{w},\boldsymbol{x})$, непрерывно дифференцируемой в области $\boldsymbol{W}\times\boldsymbol{X}$ является искомая нелинейная функция, отражающая зависимость между входным и выходным сигналом. А множество известных пар "вход-выход" $\boldsymbol{D}=\{(\boldsymbol{x_n},\boldsymbol{y_n})\}_{n=1}^{N}$, полученных на основании экспериментальных данных, представляют собой регрессионную выборку свободной и зависимой переменной. С помощью процесса обучения требуется найти такое значение вектора параметров $\boldsymbol{w}$, что доставляло бы функции ошибки локальный минимум

$$E_D=\sum_{n=1}^{N}(y_n-f(w,x_n))^2 .$$

Как и другие квазиньютоновские методы алгоритм Левенберга-Марквардта исключает необходимость явного задания матрицы Гессе. Когда функция производительности имеет вид суммы квадратов (что является типичным для обучения сетей с прямым распространением), тогда гессиан может быть рассчитан приближенно, используя выражение

$$\boldsymbol{H}=\boldsymbol{J}^T\boldsymbol{J},$$

а градиент вычислен как

$$\boldsymbol{grad}=\boldsymbol{J}^T\boldsymbol{e},$$

где $\boldsymbol{J}$ – матрица Якоби, содержащая первые производные ошибок сети относительно весов и смещений;

$\boldsymbol{e}$ – вектор ошибок сети.

В самом начале работы алгоритма задается начальный вектор параметров $\boldsymbol{w}$, который на каждом последующем шаге итерации заменяется на вектор $\boldsymbol{w}+\Delta\boldsymbol{w}$. Для оценки $\Delta\boldsymbol{w}$ пользуются линейным приближением

$$f(\boldsymbol{w}+\Delta\boldsymbol{w},\boldsymbol{x})\approx f(\boldsymbol{w},\boldsymbol{x})+\boldsymbol{J}\Delta\boldsymbol{w},$$

где $\boldsymbol{J}$ – $N \times R$-матрица Якоби функции $f(\boldsymbol{w}, \boldsymbol{x})$ в точке $\boldsymbol{w}$, которую можно представить:

$$\boldsymbol{J} = \begin{vmatrix} \dfrac{\partial f(\boldsymbol{w}, \boldsymbol{x}_1)}{\partial w_1} & \cdots & \dfrac{\partial f(\boldsymbol{w}, \boldsymbol{x}_1)}{\partial w_R} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f(\boldsymbol{w}, \boldsymbol{x}_N)}{\partial w_1} & \cdots & \dfrac{\partial f(\boldsymbol{w}, \boldsymbol{x}_N)}{\partial w_R} \end{vmatrix},$$

где $\boldsymbol{w} = [w_1, ..., w_R]^T$ – вектор параметров.

В результате, для нахождения $\Delta\boldsymbol{w}$, необходимо решить систему линейных уравнений

$$\Delta\boldsymbol{w} = (\boldsymbol{J}^T\boldsymbol{J} + \mu\boldsymbol{I})^{-1}\boldsymbol{J}^T(\boldsymbol{y} - \boldsymbol{f}(\boldsymbol{w})),$$

где μ – параметр регуляризации, который назначается на каждой итерации алгоритма,

$\boldsymbol{I}$ – единичная матрица.

Алгоритм останавливается в случае, если значение параметра приращения $\Delta\boldsymbol{w}$ в последующей итерации менее заданного значения или найденные параметры $\boldsymbol{w}$ доставляют ошибку E_D менее установленной величины. В таком случае значение $\boldsymbol{w}$ в последней итерации является искомым.

Описанные шаги повторяются до тех пор, пока не выполнится любое из условий:

- количество повторов алгоритма достигнет максимального значения;
- параметр времени превысит установленный максимум;
- эффективность примет заданное значение;
- градиент эффективности станет меньше заданной величины.

Окно с результатами обучения НС, имитирующей работу аппаратуры передающего конца и рельсовой линии показано на рис. 3.

На рис. 3 в пункте "Algorithms" описаны алгоритмы, использовавшиеся в процессе обучения. Распределение данных осуществлялось с помощью функции dividerand. Она делит все данные обучающей выборки на 3 множества: данные, используемые для обучения сети (70%), для проверки (15%) и для испытания функционирования сети (15%). Проверочное множество позволяет остановить работу алгоритма при достижении оптимальной обобщающей функции для НС, а данные для испытания ее работы позволяют провести независимую оценку точности.

Рис. 3. Окно обучения нейронной сети

В качестве целевой функции выбрано значение среднеквадратической ошибки (mean squared error – mse)

$$mse = \frac{\sum_{i=1}^{N}(y_i - f(\boldsymbol{w}, \boldsymbol{x}_i))}{N-2}.$$

Пункт "Progress" характеризует проведенный процесс обучения. На основании рис. 3 можно сделать вывод, что алгоритм завершил свою работу, потому что на тринадцатом шаге итерации значение градиента стало меньше заданной величины $1 \cdot 10^{-5}$.

С целью расширения возможностей использования схемы дополнительно введен блок, позволяющий задать сигнал помехи, воздействующий на путевой приемник РЦ. При этом появляется возможность получить на входе приемника суммарный сигнал, состоящий из полезного сигнала контроля состояния рельсовой линии и данной помехи.

Рассмотрим на примере функционирование данной НС. Исходными данными для входного сигнала будут: температура окружающей среды T=10^{o}C, сопротивление изоляции $R_{из}$=1 Ом, а также дополнительно введенная флуктуационная помеха с нулевым математическим ожиданием и единичной дисперсией (рис. 4 а).

В результате моделирования получен сигнал контроля состояния рельсовой линии на входе путевого приемника при заданном сопротивлении изоляции и температуре окружающей среды (рис. 4 б) и суммарный сигнал на входе приемника, состоящий из сигнала контроля РЦ и заданной флуктуационной помехи (рис. 4 в).

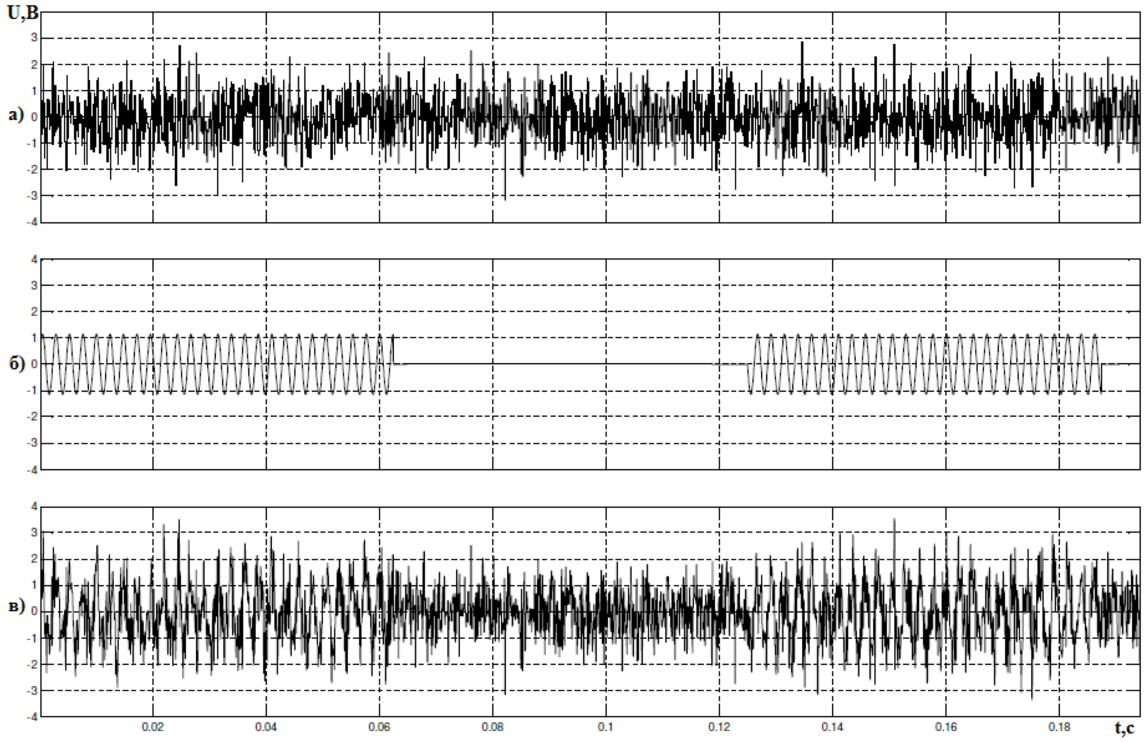

Рис. 4. Вид а) флуктуационной помехи; б)сигнала контроля состояния РЦ на выходе нейросетевой модели; в) сигнала, искаженного под воздействием помехи.

Разработанная модель имитирует работу аппаратуры РЦ в заданных условиях воздействия внешних факторов, что позволяет использовать ее в целях прогнозирования и диагностики состояния оборудования с учетом колебания параметров окружающей среды.

Выводы

На основании анализа достоинств нейронных сетей разработана нейросетевая модель рельсовой цепи. Благодаря свойству адаптивности модели учтено воздействие колебаний внешних факторов на ее работу.

Рассмотрен метод обучения нейронной сети на основании алгоритма Левенберга-Марквардта. Приведены результаты обучения и пример функционирования нейросетевой модели, показывающий возможность ее использования с целью прогнозирования и диагностики состояния аппаратуры РЦ в условиях колебания внешних факторов.

Литература

1. Хайкин С. Нейронные сети: полный курс [Текст]: [2-е изд., испр.] / С. Хайкин. – М.: ООО «И.Д. Вильямс», – 2006. – 1104 с.
2. Шулешко А.Н. Применение систем искусственного интеллекта для прогнозирования параметров качества машин [Текст] / А.Н. Шулешко, П.А. Лонцих // Вестник Иркутского государственного технического университета. – Иркутск: Иркутский государственный технический университет, 2013. –№11(82). – С.418-422.
3. Гусев К.Ю. Нейросетевое моделирование динамики нелинейных систем [Текст] / К.Ю. Гусев, В.Л. Бурковский // Вестник Воронежского государственного технического университета. – Воронеж: Воронежский государственный технический университ, 2012. – №12-1. – С.51-56.

4. Васильев А.Н. Нейросетевое моделирование в математической физике [Текст] / А.Н. Васильев // Нейрокомпьютеры: разработка, применение. – М.: "Радиотехника", 2009. – №5. – С.25-38.
5. Стрижов В.В. Методы индуктивного порождения регрессионных моделей [Текст] / В.В. Стрижов.– М.: Вычислительный центр РАН, 2008.– 61с.
6. Levenberg, K. A. Method for the Solution of Certain Problems in Last Squares [Text] / K.A. Levenberg // Quarterly of Applied Mathematics.– USA:Brown University, 1944.– Vol. 2.– P. 164—168.
7. Демиденко Е.З. Оптимизация и регрессия [Текст] / Е.З. Демиденко.– М.: Наука, 1989.– 296с.

Саяпіна І.О. Нейромережеве моделювання апаратури рейкових кіл. З метою прогнозування та діагностики стану обладнання рейкового кола розроблена нейромережева модель апаратури рейкового кола. Розглянуто метод навчання нейронної мережі на основі алгоритму Левенберга-Марквардта. Приведені результати навчання та приклад функціонування моделі.
Ключові слова: нейронна мережа, рейкове коло, нейромережеве моделювання, алгоритм Левенберга-Марквардта, навчання нейронної мережі.

Saiapina I.O. Neural network modeling of a track circuit equipment. A neural network model of the track circuit has been worked out with the purpose of track circuit equipment prediction and diagnostics. Levenberg-Marquardt neural network training algorithm has been considered. Training results and an example of model functioning have been presented.
Key words: neural network, track circuit, neural network modeling, Levenberg-Marquardt algorithm, neural network training.

Рецензент д.т.н., профессор Бабаев М.М. (УкрГАЖТ)

Поступила 10.02.2014г.