

стирования являются также и его существенным недостатком, обусловленным сроками преподавания физики в средних учебных заведениях. В школах эту проблему можно решить относительно безболезненно путем небольшого перераспределения учебной нагрузки по годам (ведь в школе физику изучают в течение пяти лет и, следовательно, больше возможностей для маневра) так, чтобы к апрелю месяцу основной курс был завершен.

В подготовительном отделе ситуация выглядит иначе из-за жестких временных ограничений. Как уже было упомянуто выше, занятия на ПО начинаются только 1 декабря, и нетрудно посчитать, что к апрелю при традиционной схеме обучения будет пройдено не более 70% от всего курса. А если учесть, что в начале марта у слушателей начинаются двухнедельные каникулы, то получится и того меньше.

Для выхода из создавшейся ситуации предлагаются следующие варианты.

1. Перенести начало учебных занятий на ПО с 1 декабря на 1 октября, а набор слушателей проводить в течение сентября. Этот вариант является наиболее простым и автоматически снимал бы большинство проблем, однако реализация его связана с техническими и организационными (зачастую надуманными) трудностями.
2. Пересмотреть структуру преподавания курса физики. Например, отказаться от традиционного метода, когда лекционный курс и практические занятия проводятся параллельно, а прочитать сначала весь теоретический материал, после чего переходить к решению тестовых задач. Оставшееся после централизованного тестирования время посвятить решению более серьезных задач для подготовки к вступительным экзаменам (если их к тому времени не

отменят). Данный вариант хоть и соблазнительно, но связан с невероятно большой нагрузкой (пять лекций в неделю), как для лектора, так и для самих слушателей, и поэтому также нереален.

3. Существует третий, компромиссный вариант. Суть его состоит в следующем. Отказаться от периодичности проведения занятий и провести их в «рваном» режиме. Лекции: декабрь – 16 часов (с учетом первой лекционной недели); январь – 18 часов (за счет некоего «вакуума», возникшего во время сессии); февраль – 16 часов; март – 8 часов (из-за каникул). Получается 58 часов при их общем объеме по плану 70 часов. Очевидно, за такое количество часов можно изложить основной материал, при этом работая в «сдающем» режиме. Что касается практических занятий (по плану – 206 часов), то их нетрудно привязать к лекционным.

Авторы далеки от мысли, что после публикации данной статьи все проблемы (или хотя бы их часть) будут решены. Мы лишь предлагаем варианты их решения и готовы взять на себя часть ответственности. Следует отметить, что системное решение таких вопросов лежит в совершенно иной плоскости. Факультет довузовской подготовки (ФДП), куда входят подготовительное отделение и подготовительные курсы (вечерние и заочные) занимает важное, специфическое место в общей системе форм непосредственной связи между средним и высшим образованием. Он удовлетворяет необходимую потребность в ликвидации «белых пятен» в школьном образовании и позволяет сделать важные выводы для сокращения дистанции между средним и высшим образованием, а в итоге, к разработке системы непрерывного образования в Республике Беларусь. Но это уже материал для другой статьи.

УДК 681.324

Савицкий Ю.В., Головкин В.А.

НЕЙРОСЕТЕВАЯ МОДЕЛЬ АНАЛИЗА ХАОТИЧЕСКИХ СИГНАЛОВ НА БАЗЕ ФУНКЦИИ АКТИВАЦИИ С НАСТРАИВАЕМЫМИ ПАРАМЕТРАМИ

1. ФУНКЦИЯ АКТИВАЦИИ С НАСТРАИВАЕМЫМИ ПАРАМЕТРАМИ

Для построения нейронной сети используется многослойная архитектура с нелинейной функцией активации в скрытом слое и линейным нейроном в выходном слое [1]. В рамках данной работы предпринята попытка использования в скрытом слое нейронной сети функции активации с настраиваемыми параметрами [2] вместо традиционных функций сигмоидального типа.

Адаптивная функция активации нейрона j определяется выражением

$$F(S_j) = a_{1j} e^{-b_{1j} S_j^2} + \frac{a_{2j}}{1 + e^{-b_{2j} S_j}} \quad (1)$$

Отличия рассматриваемой функции являются параметры a_{1j} , a_{2j} , b_{1j} , b_{2j} , представляющие собой коэффициенты, которые подлежат настройке в процессе обучения нейронной сети наряду с весовыми коэффициентами нейронных элементов. Обоснованием применения данной функции активации является доказательство [2], что соответствующая нейронная сеть является универсальным аппроксиматором для любой непрерывной функции. Особенностью рассматриваемой архитектуры нейронной сети является возможность индивидуальной

настройки параметров адаптивной функции активации для каждого нейронного элемента.

2. МОДИФИКАЦИЯ МЕТОДА ОБУЧЕНИЯ ДЛЯ ФУНКЦИИ АКТИВАЦИИ С НАСТРАИВАЕМЫМИ ПАРАМЕТРАМИ

В основу алгоритма обучения нейронной сети положен алгоритм обратного распространения ошибки, а также метод градиентного спуска для модификации весовых коэффициентов нейронных элементов и параметров адаптивной функции активации.

Весовые коэффициенты нейронной сети w_{ij} в момент времени $t+1$ изменяются в соответствии с методом градиентного спуска

$$w_{ij}(t+1) = w_{ij}(t) - \alpha \frac{\partial E}{\partial w_{ij}(t)}, \quad (2)$$

где α – шаг обучения нейронной сети, E – ошибка нейронной сети.

При использовании алгоритма обратного распространения ошибки для обучения нейронной сети производная среднеквадратической ошибки определяется как

Савицкий Юрий Викторович, к.т.н., доцент каф. ЭВМ и С Брестского государственного технического университета.

Головкин Владимир Адамович, д.т.н., зав. каф. интеллектуальных информационных технологий Брестского государственного технического университета.

Беларусь, БГТУ, 224017, г. Брест, ул. Московская, 267.

$$\frac{\partial E}{\partial w_{ij}} = \gamma_j F'(S_j) y_i. \quad (3)$$

В данной формуле γ_j – ошибка j -го нейрона определяется как $\gamma_j = y_j - t_j$ для нейронных элементов выходного слоя (t_j – эталонное значение выхода) и $\gamma_j = \sum_i \gamma_i F'(S_i) w_{ji}$ для нейронных элементов скрытого слоя (здесь индекс i характеризует нейронные элементы слоя, следующего за j); $F'(S_j)$ – производная функции активации нейронного элемента j по взвешенной сумме S_j его входов.

Производная $F'(S_j)$ определяется как

$$F'(S_j) = -2a_{1j}b_{1j}S_j e^{-b_{1j}S_j} + \frac{a_{2j}b_{2j}e^{-b_{2j}S_j}}{(1 + e^{-b_{2j}S_j})^2}. \quad (4)$$

Параметры функции активации a_{1j} , a_{2j} , b_{1j} , b_{2j} модифицируются при обучении согласно выражениям (5)-(8):

$$a_{1j}(t+1) = a_{1j}(t) - \alpha_f \frac{\partial E}{\partial a_{1j}}, \quad (5)$$

$$b_{1j}(t+1) = b_{1j}(t) - \alpha_f \frac{\partial E}{\partial b_{1j}}, \quad (6)$$

$$a_{2j}(t+1) = a_{2j}(t) - \alpha_f \frac{\partial E}{\partial a_{2j}}, \quad (7)$$

$$b_{2j}(t+1) = b_{2j}(t) - \alpha_f \frac{\partial E}{\partial b_{2j}}, \quad (8)$$

где α_f – шаг модификации параметров функции.

Найдем производные среднеквадратичной ошибки по параметрам функции активации

$$\frac{\partial E}{\partial a_{1j}} = \frac{\partial E}{\partial y_j} \frac{\partial y}{\partial a_{1j}}, \quad (9)$$

где $\frac{\partial E}{\partial y_j} = \gamma_j$, (10)

а $\frac{\partial y}{\partial a_{1j}} = F'(a_{1j}) = e^{-b_{1j}S_j^2}$. (11)

Аналогично:

$$\frac{\partial E}{\partial b_{1j}} = \frac{\partial E}{\partial y_j} \frac{\partial y}{\partial b_{1j}}, \quad \frac{\partial y}{\partial b_{1j}} = F'(b_{1j}) = -a_{1j}S_j e^{-b_{1j}S_j^2}, \quad (12)$$

$$\frac{\partial E}{\partial a_{2j}} = \frac{\partial E}{\partial y_j} \frac{\partial y}{\partial a_{2j}}, \quad \frac{\partial y}{\partial a_{2j}} = F'(a_{2j}) = \frac{1}{1 + e^{-b_{2j}S_j}}, \quad (13)$$

$$\frac{\partial E}{\partial b_{2j}} = \frac{\partial E}{\partial y_j} \frac{\partial y}{\partial b_{2j}}, \quad \frac{\partial y}{\partial b_{2j}} = F'(b_{2j}) = \frac{a_{2j}S_j e^{-b_{2j}S_j}}{(1 + e^{-b_{2j}S_j})^2}. \quad (14)$$

Таким образом, получены конечные выражения для обучения нейронной сети, использующей функцию активации вида (1):

$$w_{ij}(t+1) = w_{ij}(t) - \alpha_f \gamma_j \left(-2a_{1j}b_{1j}S_j e^{-b_{1j}S_j^2} + \frac{a_{2j}b_{2j}e^{-b_{2j}S_j}}{(1 + e^{-b_{2j}S_j})^2} \right) y_i, \quad (15)$$

$$a_{1j}(t+1) = a_{1j}(t) - \alpha_f \gamma_j e^{-b_{1j}S_j^2}, \quad (16)$$

$$b_{1j}(t+1) = b_{1j}(t) + \alpha_f \gamma_j a_{1j} S_j e^{-b_{1j}S_j^2}, \quad (17)$$

$$a_{2j}(t+1) = a_{2j}(t) - \alpha_f \gamma_j \frac{1}{1 + e^{-b_{2j}S_j}}, \quad (18)$$

$$b_{2j}(t+1) = b_{2j}(t) - \alpha_f \gamma_j \frac{a_{2j}S_j e^{-b_{2j}S_j}}{(1 + e^{-b_{2j}S_j})^2}. \quad (19)$$

3. МОДИФИЦИРОВАННЫЙ АЛГОРИТМ ОБРАТНОГО РАСПРОСТРАНЕНИЯ ОШИБКИ

Особенностью процедуры обучения рассматриваемой архитектуры нейронной сети является необходимость на каждой итерации обучения как модификации синаптических связей нейронов, так и модификации настраиваемых параметров функций активации нейронов.

1. Задается желаемая среднеквадратическая ошибка нейронной сети и шаг обучения α , α_f (возможно использование адаптивного шага обучения [3]). Инициализируются случайным образом весовые коэффициенты нейронной сети (w_{ij}) и параметры адаптивной функции активации (a_{1j} , a_{2j} , b_{1j} , b_{2j}).

2. Для каждого из последовательно подаваемых на вход образов обучающей выборки выполняются следующие действия:

2.1. Производится фаза прямого распространения информации по нейронной сети. Вычисляется выходная активность всех нейронных элементов сети y_j .

2.2. Осуществляется фаза обратного распространения информации, в результате которой определяется ошибка γ_j для всех нейронных элементов каждого слоя нейронной сети.

2.3. Производится изменение весовых коэффициентов нейронных элементов для каждого слоя нейронной сети. При этом для слоя нейронной сети, содержащего нейроны с настраиваемой функцией активации, весовые коэффициенты модифицируются по правилу (15).

2.4. Выполняется модификация параметров адаптивной функции активации для всех нейронных элементов адаптивного слоя согласно правил (16) – (19).

3. Вычисляется суммарная среднеквадратическая ошибка нейронной сети

$$E = \frac{1}{2} \sum_{k=1}^L \sum_j (y_j^k - t_j^k)^2, \quad (20)$$

где L – размерность обучающей выборки.

Шаг 2 алгоритма должен быть выполнен многократно до достижения заданной точности нейросетевой модели.

4. АНАЛИЗ И ПРОГНОЗИРОВАНИЕ ХАОТИЧЕСКИХ СИГНАЛОВ

Для экспериментальных исследований рассматриваемого метода использовалась нейронная сеть с 2 входными элементами, 5 скрытыми и 1 выходным нейроном для предсказания хаотических процессов Лоренца и Энона [1]. Как известно, хаотический процесс Энона задается выражением

$$\begin{cases} x_{n+1} = 1 - \alpha x_n^2 + y_n \\ y_{n+1} = \beta x_n \end{cases}, \quad (19)$$

где $\alpha = 1,4$ и $\beta = 0,3$ для хаотического поведения системы.

Аттрактор Лоренца описывается системой трех дифференциальных уравнений:

$$\begin{cases} \frac{dx}{dt} = G(y - x) \\ \frac{dy}{dt} = -xz + rx - y \\ \frac{dz}{dt} = xy - bz \end{cases} \quad (20)$$

Процесс Лоренца является хаотической системой для значений параметров $G=10$, $r=28$ и $b=8/3$.

Элементы скрытого слоя сети имеют адаптивную функцию активации, выходной нейрон – линейную. Использование функции активации с настраиваемыми параметрами в скрытом слое нейронной сети описанной конфигурации позволяет осуществлять прогнозирование с точностью сравнимой с получаемой при использовании сигмоидальной функции активации в скрытом слое сети конфигурации 7-5-1 [1]. При этом время, затрачиваемое на обучение нейронной сети с

адаптивной функцией активации в скрытом слое, оказывается значительно меньше за счет оптимизации архитектуры сети. Обучающая выборка состоит из 50 элементов для системы Энона и 400 элементов для системы Лоренца соответственно. Шаг обучения был выбран равным 0.01 и 0.005 для скрытого и выходного слоя нейронной сети при обучении нейронной сети на данных ряда Энона и Лоренца соответственно. Шаг изменения параметров адаптивной функции активации был выбран соответственно равным 0,01 и 0,001. Среднеквадратичная ошибка обучения сети составила $7,93 \cdot 10^{-6}$ после 500 итераций для данных ряда Энона и $9,84 \cdot 10^{-4}$ после 500 итераций для данных ряда Лоренца. Следует отметить, что аналогичные эксперименты, проведенные на основе нейронной сети 7-5-1 с сигмоидной функцией активации, дали следующие результаты [1]:

- для процесса Энона: число итераций обучения – 3000, достигнутая погрешность обучения - $3,3 \cdot 10^{-4}$;
- для процесса Лоренца: число итераций обучения – 3000, достигнутая погрешность обучения - $8,1 \cdot 10^{-4}$.

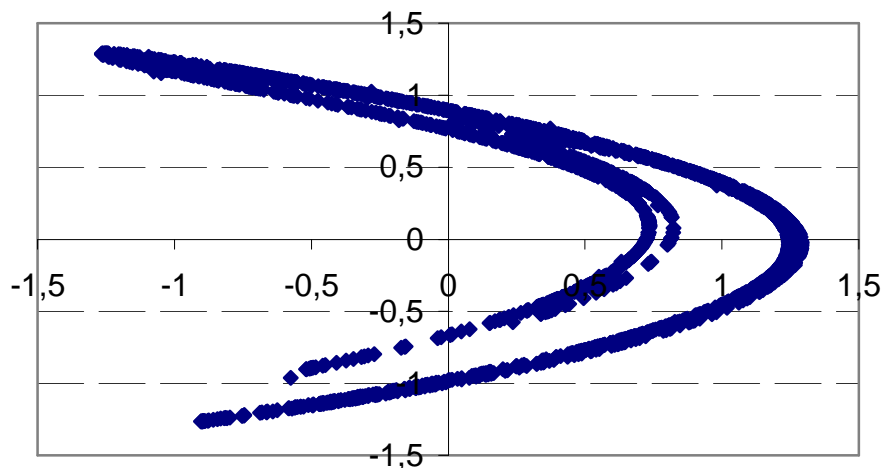


Рис. 1. Аттрактор системы Энона, построенный по результатам прогнозирования

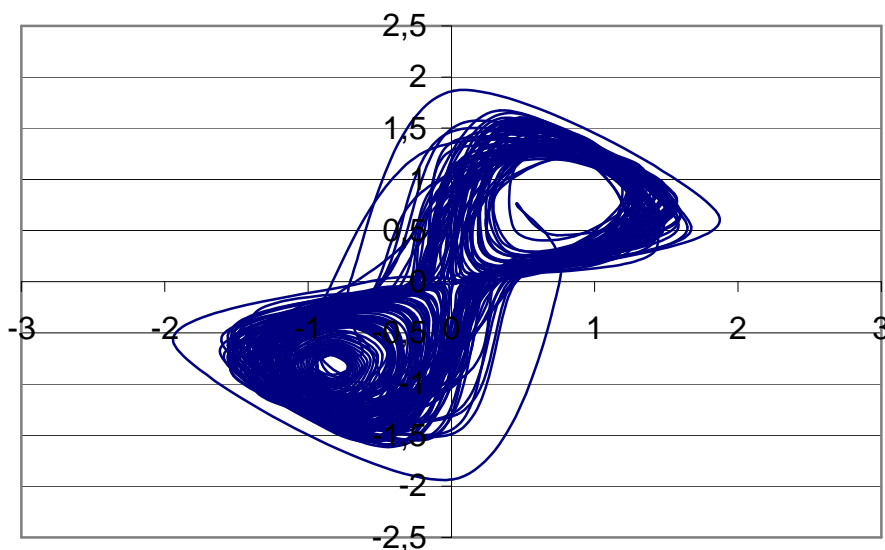


Рис. 2. Аттрактор системы Лоренца, построенный по результатам прогнозирования

По результатам прогнозирования, проведенного при помощи нейронной сети, построены аттракторы систем Энона и Лоренца. Объем прогноза для ряда Энона – 2000 значений, для ряда Лоренца – 500 значений. На рис. 1 и рис. 2 представлены соответствующие аттракторы систем Энона и Лоренца, построенные по результатам многошагового прогноза.

По разработанной авторами данной работы методике [1] для описанных хаотических процессов на основе полученных нейросетевых моделей были рассчитаны старшие показатели Ляпунова: для процесса Энона $\lambda = 4,42$; для процесса Лоренца $\lambda = 0,963$. Полученные значения позволяют сделать вывод о пригодности рассматриваемых в работе нейросетевых моделей для решения круга задач, связанных с анализом и обработкой хаотических сигналов.

Таким образом, нейронная сеть, построенная с использованием адаптивной функции активации в скрытом слое, имеет высокую скорость обучения и достаточно малую ошибку при обучении и прогнозировании. Использование адаптивной функции активации позволяет существенно уменьшить количество нейронных элементов в скрытом слое за счет индивидуальной подстройки параметров функций активации для каждого из нейронов скрытого слоя.

Следует, однако, отметить некоторые особенности метода обучения. Поскольку речь идет об одновременной настройке в процессе обучения, как весовых коэффициентов, так и параметров функции активации, то метод является требователь-

ным к выбору размеров шагов обучения α, α_f (а также их соотношения) в рамках отдельной архитектуры сети и обучающей выборки.

ЗАКЛЮЧЕНИЕ

В данной статье описан метод применения функции активации с настраиваемыми параметрами при построении нейронных сетей. Принципиальное отличие предложенного метода – модификация параметров функции активации нейронных элементов наряду с их весовыми коэффициентами в процессе обучения нейронной сети. Изложены математические основы предлагаемого метода и результаты использования нейронной сети с адаптивной функцией активации для прогнозирования временных рядов.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. V. Golovko, Y. Savitsky, N. Maniakov. Neural Networks for Signal Processing in Measurement Analysis and Industrial Applications: the Case of Chaotic Signal Processing // chapter of NATO book "Neural networks for instrumentation, measurement and related industrial applications". - Amsterdam: IOS Press, 2003, pp. 119-143.
2. Shuxiang Xu, Ming Zhang «Justification of a neuron-adaptive activation function». Proceedings of IEEE 2000.
3. Голловко В.А. Нейроинтеллект: теория и применение. Организация и обучение нейронных сетей с прямыми и обратными связями. – Брест, Изд. БПИ. – 1999. – 264с.

УДК 681.324

Головко В.А., Каменда Д.В., Кочурко П.А.

НЕКОТОРЫЕ АСПЕКТЫ ПРИМЕНЕНИЯ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ОБНАРУЖЕНИЯ АТАК

1. ВВЕДЕНИЕ

Стремительный и всесторонний рост Интернет-технологий увеличивает важность защиты компьютерных сетей от атак. За последние годы число сетевых атак очень быстро возросло, что привело к значительным проблемам в различных компаниях. К примеру, DoS-атаки ("Denial of Service" – отказ в обслуживании) на такие компании, как Yahoo, принесли им убытки в миллионы долларов.

Системы обнаружения атак (СОА) используются как средство обеспечения безопасности компьютерных сетей и выдают администратору сигнал тревоги в случае атаки. Основная цель СОА – обнаружение и распознавание сетевых атак в режиме реального времени. Сейчас существуют различные подходы к обнаружению атак. Это сигнатурный метод, экспертные системы, встроенные сенсоры, нейронные сети, искусственные иммунные системы [1-6] и т. д. Большинство таких СОА могут обнаруживать известные атаки и имеют очень мало возможностей для обнаружения новых атак.

Данная статья описывает применение нейронных сетей для обнаружения атак путем анализа данных сетевого трафика. Это базируется на том, что отказ в обслуживании и другие сетевые атаки представлены в сетевом трафике. Поэтому использование нейронных сетей позволяет выделить нелинейные зависимости между переменными из сетевого трафика и проецировать системы обнаружения атак реального времени.

В статье описывается система обнаружения атак, которая состоит из двух различных нейронных сетей. Первой нейронной сетью является нелинейная рециркуляционная нейронная сеть (РНС), которая позволяет идентифицировать нормальное или аномальное поведение системы. Вторая сеть – многослойный перцептрон (MLP), который может распознавать тип атаки.

Статья организована следующим образом. В разделе 2 описана система обнаружения атак, основанная на нейросетевом подходе. Часть 3 описывает нелинейную РНС и многослойный перцептрон для идентификации и классификации компьютерных сетевых атак. В части 4 представлены результаты экспериментов. Выводы даны в части 5.

2. ОПИСАНИЕ СИСТЕМЫ

Рассмотрим блок-схему системы обнаружения атак (рисунок 1). Она состоит из нескольких этапов. В начале система считывает данные сетевого трафика, которые поступают в модуль предобработки. Задача модуля предобработки – сбор необходимых данных для нейронных сетей из сетевого трафика.

Функционирование его базируется на WinPcap, в результате чего из сетевого трафика выделяются элементы, показанные в таблице 1, которые используются в дальнейшем для обучения и тестирования нейронных сетей. Соединение –

Каменда Дмитрий Васильевич, студент 5-го курса Брестского государственного технического университета.

Кочурко Павел Анатольевич, аспирант каф. интеллектуальные информационные технологии Брестского государственного технического университета.

Беларусь, БГТУ, 224017, г. Брест, ул. Московская, 267.

Физика, математика, информатика