

Ж. Ш. Шаршеналиев, avtomatika_nankr@mail.ru
Институт машиноведения и автоматизации НАН КР,
Ш.А. Мирзакулова, mirzakulova@mail.ru,
С.У. Исакова., Sabira49@mail.ru

ПРОГНОЗИРОВАНИЕ СЕТЕВОГО ТРАФИКА

Современная гетерогенная сеть порождает сетевой трафик со сложной (неоднородной) структурой. Исследование реально измеренных данных показывают, что они обладают не равномерной интенсивностью поступления пакетов на обслуживаемые сетевые устройства. При этом модели, построенные на основе данных, характеризующего один объект за ряд последовательных моментов времени (временные ряды) обладают свойством нестационарности, что означает, что их структура многокомпонентная. Поэтому сегодня задачи управления сетью основываются, в том числе и на предсказанных будущих данных для принятия правильного решения. Выявить и количественно выразить компоненты сложной структуры – наличие/отсутствие тенденции, периодичности, случайной компоненты является основной задачей анализа временного ряда. Идентифицировать нелинейную функцию и осуществить его прогнозирование с большим успехом сегодня осуществляют алгоритмы нейронной сети с глубоким обучением.

Ключевые слова: временной ряд, нейронная сеть, перцептрон, многослойная сеть, прогноз

Введение

Проведенные исследования временного ряда сгенерированного в мульти-сервисной пакетной сети показали неравномерную интенсивность количества пакетов, то есть ряд оказался нестационарным. Для прогнозирования таких рядов классические методы прогнозирования не пригодны.

Нелинейную идентификацию с прогнозированием возможно с использованием нейросетевого прогнозирования. При этом нейронные сети и глубокое обучение хорошо справляются с подобными задачами.

Еще в 1943 г. была написана статья нейрофизиологом Уорреном МакКаллохом и математиком Уолтером Питтсом о том, как могут работать нейроны в мозге. Они показали, что с помощью пороговых нейронных элементов можно реализовать исчисление логических функций для распознавания образов. Была смоделирована простая нейронная сеть с использованием электрической цепи. В 1949 г. Дональдом Хеббом было предложено правило обучения. При этом было отмечено, что нейронные пути усиливаются при каждом их использовании. По его словам, если два нерва срабатывают одновременно, связь между ними усиливается. В 1950 г. стало возможным на компьютерах осуществлять моделирование гипотетической нейронной сети (сотрудник лаборатории ИВМ Натаниал Рочестер). Далее были разработаны распознавание двоичных шаблонов, адаптивный фильтр, устраняющий эхо на телефонных линиях. В 1962 г. Фрэнк Розенблатт предложил свою модель нейронной сети – перцептрон, что послужило крупным толчком в развитии нейрокибернетики. Эта модель основана на идее, что хотя бы один активный перцептрон может иметь большую ошибку, то можно отрегулировать значения веса с целью распределения его по сети или, по крайней мере, для соседних перцептронов и получить предсказанное значение. В 1962 г. Widrow & Hoff разработали процедуру обучения, а в 1975 г. создана первая многослойная сеть, далее разработана сеть Хопфилда и т.д. В 80-е XX века

появляются термины «нейрокибернетика», «нейроинформатика», «нейрокомпьютеры». Именно доктор Роберт Хехт-Нильсен дал определение нейронной сети, которую называл искусственной нейронной сетью (ИНС) и изобрел один из первых нейрокомпьютеров. Следовательно, нейронная сеть представляет собой серию алгоритмов, которые стремятся распознать основные взаимосвязи в наборе данных посредством процесса, который имитирует работу человеческого мозга. В этом смысле нейронные сети относятся к системам нейронов, как органических, так и искусственных по своей природе. Нейронные сети могут адаптироваться к изменению ввода; таким образом, сеть генерирует наилучший возможный результат без необходимости пересмотра критериев вывода. Концепция нейронных сетей, которая имеет свои корни в искусственном интеллекте, быстро набирает популярность и используется в различных областях.

Выбор метода прогнозирования нестационарного ряда и программной оболочки

К основным методам прогнозирования нестационарного временного ряда относятся:

- статистические методы;
- новые методы на основе искусственного интеллекта (ИИ).

Среди статистических подходов метод ARIMA (Auto-Regressive Integrated Moving Average) позволяет описывать нестационарные временные ряды, которые сводятся к стационарным рядам взятием разностей некоторого порядка от исходного временного ряда.

К новым методам на основе ИИ относятся:

- ИНС – искусственные нейронные сети;
- ANFIS (Adaptive Neuro-Fuzzy Inference System) – адаптивная сеть на основе системы нечеткого вывода;
- Fuzzy Logic – нечеткая логика;
- SVM (Support Vector Machine) – метод опорных векторов;
- Neuro-Fuzzy Network - нейро-нечеткие сети;
- Evolutionary Optimization Algorithm - эволюционные алгоритмы оптимизации.

Нейронные сети состоят из простых элементов функционирующих параллельно. Название этих элементов пришло из биологических нервных систем. Как и в природе, связи между элементами четко определены функцией сети.

Возможности нейронных искусственных сетей успешно применяются для решения задач классификации, прогнозирования и управления, обеспечивая предельное распараллеливание алгоритмов, соответствующих нейросетевой технологии обработки данных.

Алгоритм обратного распространения (Back propagation) – это контролируемый алгоритм обучения, предназначенный для обучения многослойных перцептронов (ИНС), который способен создавать полезные новые функции, что отличает этот алгоритм от более ранних, более простых алгоритмов. При этом суть многослойности перцептрона заключается в том, что метод обратного распространения ошибки учит все слои, а не, потому что перцептрон имеет несколько слоев. Обратное распространение – это сущность обучения нейронной сети. Это практика точной настройки весов нейронной сети на основе частоты ошибок (то есть потерь), полученных в предыдущей эпохе (то есть итерации). Правильная настройка весов обеспечивает более низкий уровень ошибок, что делает модель надежной за счет увеличения ее обобщения.

Под обучением понимают адаптацию сети к предъявляемым эталонным образцам с помощью изменения весовых коэффициентов связей между нейронами (рис. 1).

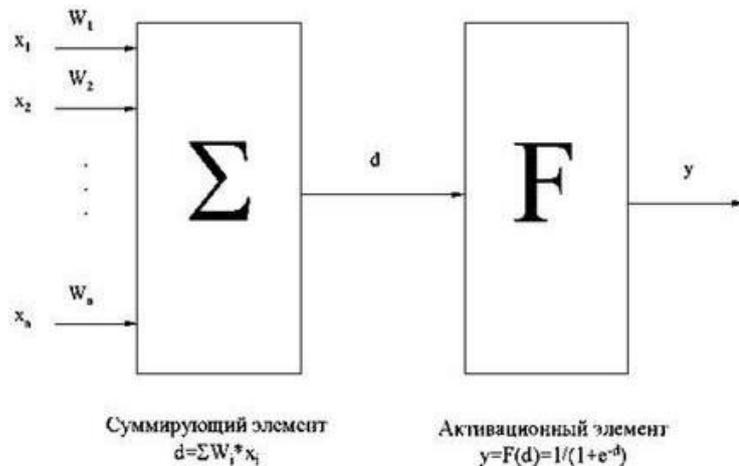


Рисунок 1 – Искусственный нейрон

Послойная организация нейронов копирует слоистые структуры определенных отделов мозга [1]. Многослойная нейронная сеть способна осуществлять любое отражение входных векторов в выходные.

Обучение нейронной сети может вестись с учителем (представляются значения, как входных данных, так и желательных выходных сигналов, и она по некоторому алгоритму осуществляется подстройка веса своих синаптических связей) и без учителя (веса изменяются по алгоритму, учитывающие только входные и производные от них сигналы). Нейронная сеть, обучаемая с учителем – многослойная и пригодна для решения задачи нелинейной авторегрессии.

Для осуществления прогнозирования нестационарного временного ряда выбран пакет прикладных программ (ППП) фирмы Math Works Neural Network Toolbox (NNT), который функционирует под управлением ядра системы Matlab.

В пакете NNT встроены следующие обучающиеся функции:

- trainidx – МОРО с импульсом и адаптацией;
- trainedf – сопряженных градиентов (Поллака-Рибейры);
- trainlm – метод Левенберга-Марквардта.

Точность прогнозирования зависит от ее архитектуры, то есть от числа скрытых слоев и числа нейронов в каждом слое. У перцептрона, который имеет много слоев вычислительная способность более высока и он хорошо описывают многомерные зависимости с большой степенью нелинейности, чем перцептрон с одним слоем.

К преимуществам ППП NNT можно отнести – большое количество различных моделей и методов прогнозирования временного ряда, создания и использования многослойных перцептронов и др.

Прогнозирование нестационарного ряда в программной среде Matlab

Измеренный ряд динамики приведен на рисунке 2, который отображает совокупность переданных по магистральной сети пакетов за пять часов за каждые 10 с.

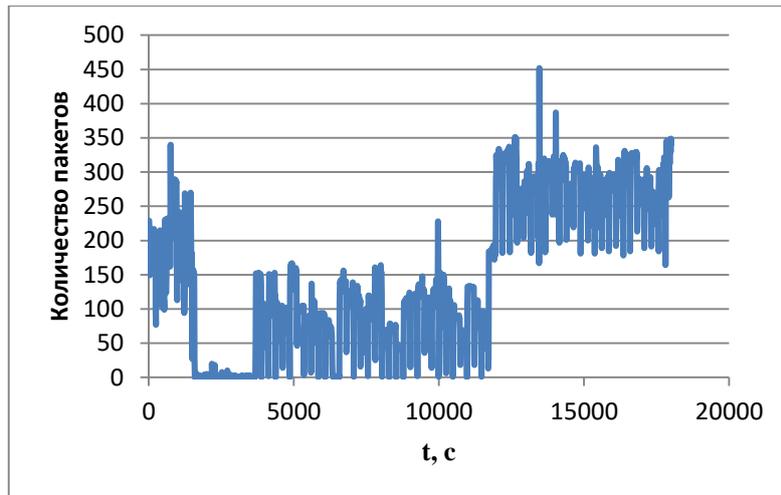


Рисунок 2 – Ряд интенсивности передачи пакетов

Для осуществления прогнозирования сетевого трафика воспользуемся ИНС. Для обучения был представлен временной ряд интенсивности передачи пакетов. Выходные – прогнозные данные. Была использована сеть с пятью слоями, в том числе в ней имеется четыре скрытых слоя. Первый скрытый слой связан со вторым, второй связан с третьим и третий – с четвертым. Первый скрытый слой содержит 50 нейронов, второй – 20, третий – 20, четвертый – 20, пятый – 1 (рис. 3).

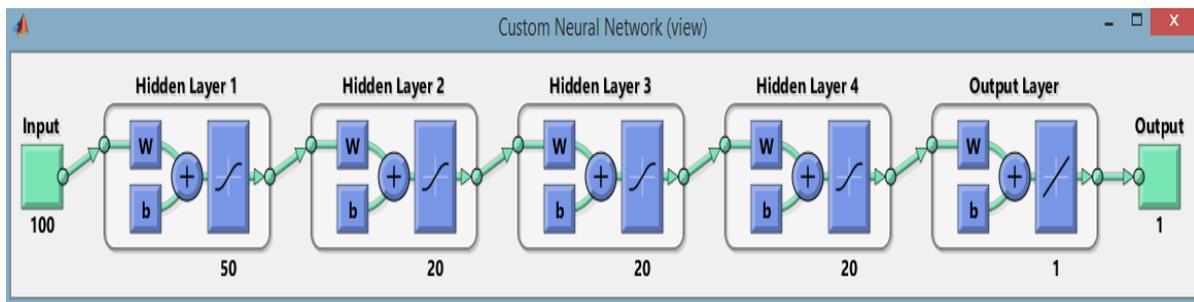


Рисунок 3 – Структура ИНС

В результате запуска и обучения ИНС получены информативные графики:

- Train, описывающий параметры настройки обучения. При этом значение эпохи (цикл) равно 1, что показывает сеть, представлена вводом нового шаблона корректировки (обратное распространение ошибки) вкупе с прямым потоком активации произошли под корректировки весов ИНС с итоговой среднеквадратичной погрешностью. Дальнейшее увеличение значения эпохи приводит к уменьшению значения средней квадратичной ошибки MSE, то есть тренировка продолжалась 4 эпохи;

- Validation – исследуемые данные после значения эпохи равного 1 не изменяет своих значений, то есть лучшая производительность берется из эпохи с наименьшей ошибкой валидации (точка обведенная кружочком);

- Test – показывает наилучшую производительность обучения;

- Best - лучшая производительность.

График обучения представлен на рис. 4, на котором видно, что наименьшая ошибка валидации равна 0,0053042.

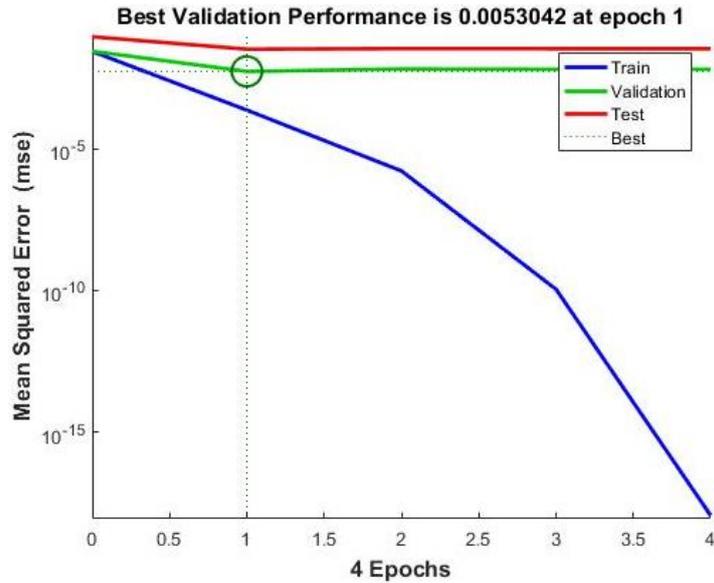


Рисунок 4 – График обучения

На рис. 5 показаны инструменты обучения ИНС, дополнительно отображающие различные параметры прогресса прогнозирования.

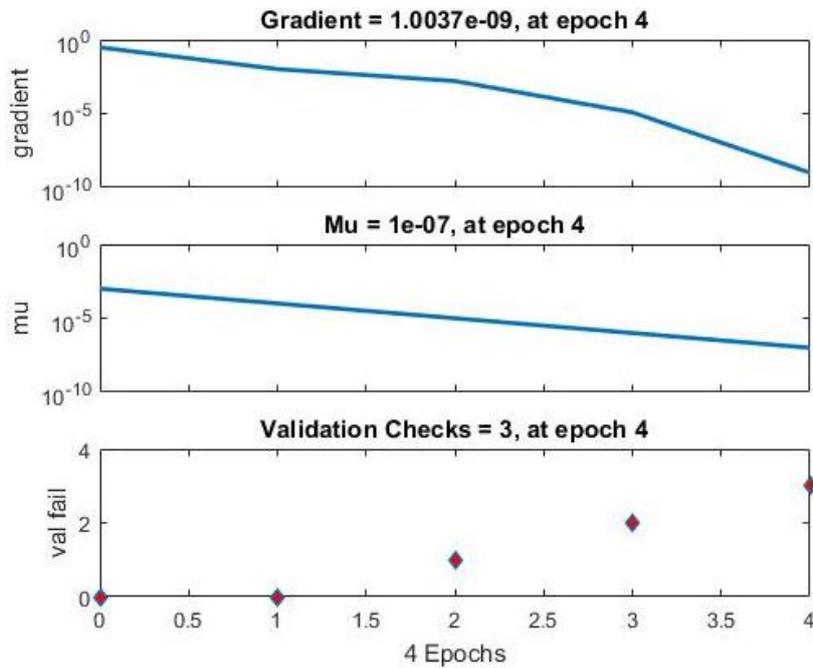


Рисунок 4 – Инструменты обучения ИНС

На рис. 5 показаны два графика данные обучения и проверки, на которых пунктирная линия на каждом графике представляет идеальный результат. Сплошная линия на графике представляет собой наилучшую линейную регрессию между результатами и целями. Значение R является показателем взаимосвязи между выходами и целями. Результаты обучения и испытаний показывают большие значения параметра R.

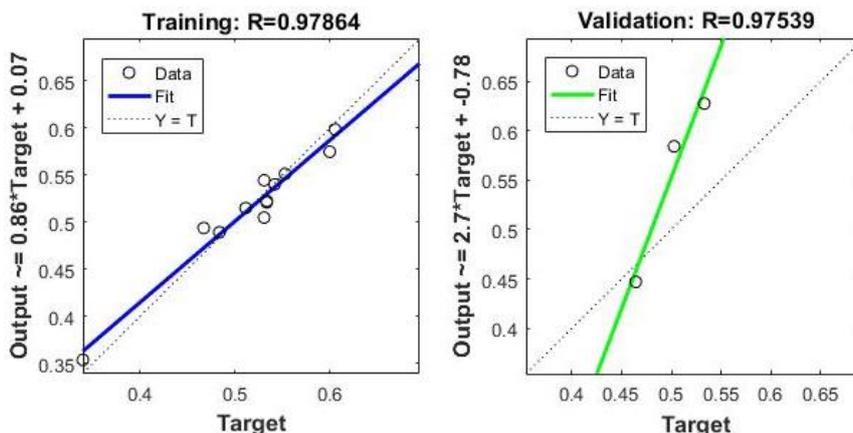


Рисунок 5 – Графики аппроксимации обучения и проверки

На рис. 6 приведен график прогноза интенсивности пакетов до уровня 2300.

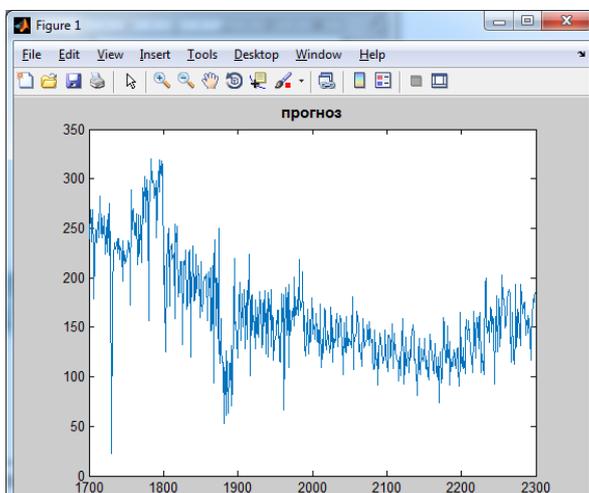


Рисунок 6 – Прогноз интенсивности пакетов

Заключение. Задача прогнозирования интенсивности трафика была решена при помощи ИНС. При этом получены: наименьшая ошибка валидации, равная 0,0053042, наилучшая аппроксимация ряда с высокими значениями параметра R.

Литература

1. Николаева С.Г. Нейронные сети. Реализация в Matlab. Учебное пособие. Нейронные сети. Реализация в Matlab: учебное пособие / С.Г. Николаева. – Казань: Казанский государственный энергетический университет, 2015. – 92 с.
2. Исакова С.У., Мирзакулова Ш.А. Статистическая оценка основных тенденций измеренного временного ряда. – Вестник КРСУ, 2018. – Том 18. – №12. –С. 51–54.
3. Шаршеналиев Ж.Ш., Мирзакулова Ш.А., Юсупова Г.М. Оценка спектра мощности временного ряда // Материалы научной конференции ИИВТ КН МОН РК «Инновационные IT и Smart-технологии» – Алматы, 03. 2019. – С. 332–337.
4. Сериков Т.Ф., Мирзакулова Ш.А., Юсупова Г.М., Авелбекова С.Ш., Сабитова А.Ж. Анализ временного ряда методом ССА // Вестник ПГУ, 2019. – № 3 – С. 328–339.