

Жуковицкий И.В., Пахомова В.Н., Дмитриев С.Ю.

Днепропетровский национальный университет железнодорожного транспорта имени академика В. Лазаряна, Украина

Повышение эффективности прогнозирования на железнодорожном транспорте

Підвищення ефективності прогнозування на залізничному транспорті

Increase of Prediction Effectiveness on Railway Transport

Представлен обзор методов прогнозирования, которые могут быть использованы на железнодорожном транспорте. Для прогнозирования простоя грузовых поездов на станции Пятихатки предлагается использовать адаптивную сеть нечеткого вывода ANFIS. Описана реализация такого прогноза в пакете расширения Fuzzy Logic Toolbox системы MATLAB.

Ключевые слова: прогнозирование, грузовые поезда, нейронечеткая сеть, обучающая выборка

В настоящее время железнодорожный транспорт сохраняет ведущее положение в общетранспортном балансе грузовых перевозок нашей страны. Однако усиление конкуренции между видами транспорта может привести к потере железными дорогами части рынка и к ухудшению их финансово-экономического положения.

Вполне справедливым является утверждение большинства ученых – экономистов, о том, что высокая инвестиционная составляющая по транспорту оправдывается только при эффективном его использовании. Ключевым моментом при анализе и планировании перевозок является возможность с максимальной точностью предсказывать значения производственных показателей, характеризующих деятельность компании. Иначе говоря, задача адекватного прогнозирования приобретает решающую значимость.

К счастью, сегодня существует масса методов построения прогнозов с высокой точностью.

Различают два основных вида прогнозирования: экспертное и формализованное.

Экспертное прогнозирование подразумевает формирование будущих значений экспертом, т.е. человеком, обладающим глубокими знаниями в определенной области. Эксперт при этом часто использует математический аппарат, однако в данном виде прогнозирования математический аппарат является лишь вспомогательным вычислительным инструментом. Экспертное прогнозирование применяется тогда, когда объект прогнозирования либо слишком прост, либо, напротив, настолько сложен, что аналитически учесть влияние внешних факторов невозможно. К таким методам относятся следующие методы: экспертных оценок, исторических аналогий, предвидения по образцу, нечеткая логика, сценарное моделирование «что – если».

Формализованное прогнозирование — это прогнозирование на основании математической модели, которая, улавливая закономерности процесса, на

своим выходе имеет будущие значения исследуемого процесса. Математических моделей прогнозирования довольно много, например, согласно ряду обзоров в настоящее время насчитывается свыше 100 классов моделей прогнозирования. Число общих классов моделей, которые в тех или иных вариациях повторяются в других, конечно, гораздо меньше и сводится к дюжине: регрессионные модели (regression model), авторегрессионные модели (auto regressive model, AR), нейросетевые модели (artificial neural network, ANN), модели экспоненциального сглаживания (exponential smoothing, ES), модели на базе цепей Маркова (Markov chain), классификационно-регрессионные деревья (classification and regression trees, CART), метод опорных векторов (support vector machine, SVM), генетический алгоритм (genetic algorithm, GA), модель на основе передаточных функций (transfer function, TF), формализованная нечеткая логика (fuzzy logic, FL), фундаментальные модели.

Ни одна из вышеперечисленных формализованных моделей прогнозирования, несмотря на ряд достоинств, не может дать абсолютно точного прогноза. Поэтому возникает необходимость в поиске и внедрении методов повышения точности прогноза. В настоящее время существуют два проверенных пути повышения точности прогнозирования. Первый путь состоит в последовательном применении двух моделей, второй путь — в параллельном использовании двух моделей.

Способ первый: кластеризация и последующее прогнозирование. Применение кластеризации, т.е. разбиение временного ряда на гомогенные отрезки, в задачах прогнозирования было предложено давно. Так, в [1] предложено прогнозирование временных рядов разбить на два этапа: кластеризация и затем прогнозирование внутри кластера

По сути дела такой подход является комбинацией двух различных моделей временных рядов – первая модель определяет кластер, вторая прогнозирует внутри кластера. Этот подход чрезвычайно популярен.

Способ второй: консенсус-прогноз. В [2] утверждается, что прогноз, полученный как комбинация нескольких независимых прогнозов, позволяет повысить точность прогнозирования.

При формировании консенсус-прогноза в расчет принимаются два и более прогноза, выполняемых независимыми организациями или моделями. Доказано [3], что точность консенсус-прогноза может быть выше точности каждого из прогнозов, принимаемых во внимание.

В работе [4] для прогнозирования времени простоя грузовых поездов станции «Пятихатки» использовалась нейронечеткая система ANFIS, которая использовала консенсус-модель из нейронной сети и нечеткой логики.

Нечеткие нейронные сети или гибридные сети призваны объединить в себе достоинства нейронных сетей и систем нечеткого вывода. Они позволяют разрабатывать и представлять модели систем в форме правил нечетких продукций, а для построения правил нечетких продукций используются возможности нейронных сетей. ANFIS (Adaptive-Network-Based Fuzzy Inference System) - адаптивная сеть нечеткого вывода, предложенная Янгом (Jang), реализованная в пакете расширения Fuzzy Logic Toolbox (пакет

нечеткой логики) системы MATLAB. ANFIS является одним из первых вариантов гибридных нейронечетких сетей, архитектура которой изоморфна нечеткой базе знаний. В нейронечетких сетях используются дифференциальные реализации треугольных норм, а также гладкие функции принадлежности [5]. Системы с нечеткой логикой целесообразно применять для сложных процессов, когда отсутствует простая математическая модель; если экспертные знания об объекте или процессе можно сформулировать только в лингвистической форме. Для решения прикладных задач наиболее часто используются треугольные, трапециевидные и «колоколообразные» функции принадлежности. С одной стороны, гибридная сеть представляет собой многослойную нейронную сеть специальной структуры без обратных связей. При этом значения входов, выходов и весов гибридной нейронной сети представляют собой действительные числа из отрезка $[0,1]$. Существующая выборка данных используется для определения параметров функций принадлежности, которые лучше всего соответствуют некоторой системе нечеткого вывода. При этом для нахождения параметров функций принадлежности используются известные процедуры обучения нейронных сетей. С другой стороны, гибридная сеть представляет собой систему нечеткого вывода FIS типа Сугено нулевого или первого порядка, в которой каждое из правил нечетких продукций имеет постоянный вес, равный 1.

Процесс создания адаптивной нейронечеткой системы можно условно разделить на следующие этапы: обработка входных данных, разработка системы, проверка системы. В качестве входных данных использовались данные о времени поступления и интервалы простоя грузовых поездов на станцию «Пятихатки» за определенный период.

Прежде чем вводить данные в систему было необходимо узнать сколько предыдущих значений понадобится для выявления закономерности появления чрезмерного простоя. Известно, что чем больше предварительных данных получит система, тем точнее будет прогноз, но обратной стороной медали является чрезмерная сложность системы, что приводит к значительному увеличению времени обработки данных и даже к зависанию системы. После ряда попыток было обнаружено, что минимальным числом входов для нахождения необходимых закономерностей простоев поездов на станции «Пятихатки» является число 6. То есть системе нужно не менее 6 входов для того, чтобы предусмотреть недопустимо долгий простой, который непременно осуществляется после 6-и значений минимального простоя. Общая структура 6-входовой нечеткой нейронной сети приведена на рис. 1.

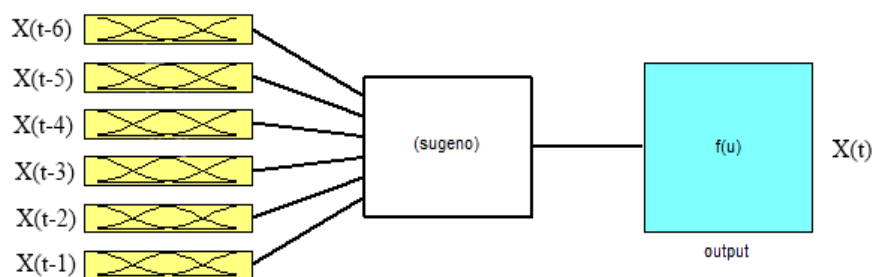


Рисунок 1 – Общая структура 6-входовой нечеткой нейронной сети

После формирования обучающей выборки, система становится нечувствительной для простоев, которые превышают 60 минут. С другой стороны, при жертвовании чувствительностью системы четко выделяются те значения, на которые надо обратить внимание при получении прогноза (60); появляется возможность интерпретировать полученные значения в бинарной системе: «допустимый простой» или «недопустимо большой простой».

Последовательность создания нечеткой нейронной сети ANFIS, предназначенной для прогнозирования периодов простоя поездов в среде Matlab следующая: запуск ANFIS-редактора; загрузка обучающей выборки; создание исходной системы нечеткого логического вывода в области генерации; выполнение обучения сети в области обучения; тестирование нечеткой системы с выводом результатов в область визуализации.

Внутренняя структура созданной прогнозирующей системы показана на рис. 2.

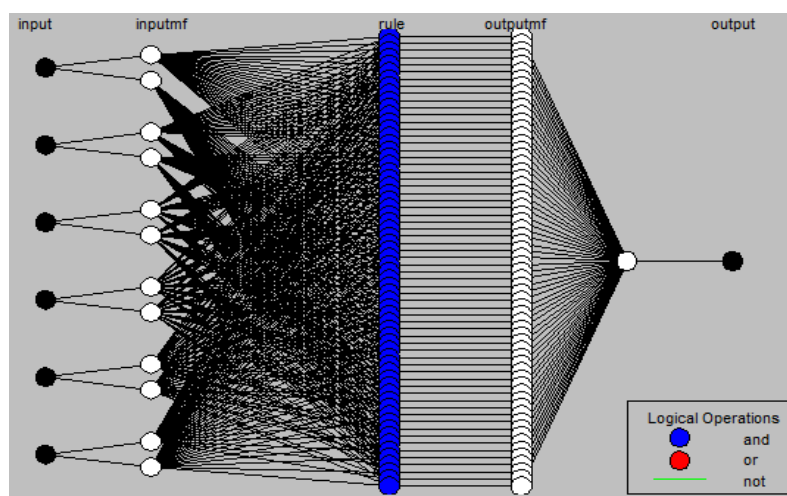


Рисунок 2 – Структура спроектированной системы нечеткого вывода ANFIS

Для проверки адекватной работы системы взято 6 последних значений обучающей выборки и использовано для получения прогноза. На базе полученных данных построены графики прогнозов периодов простоя поездов, которые представлены на рис. 3.

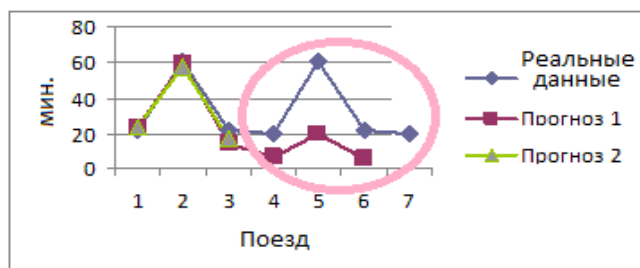


Рисунок 3 – Прогноз периодов простоя поездов

Из рис. 3 видно, что система выдает прогноз, близкий к реальным значениям только первые три шага. Потом погрешность начинает резко возрастать, и прогноз скорее отражает общую тенденцию нежели показывает реальную картину. Улучшить способность системы к прогнозированию

возможно увеличив число входов системы (это расширит сферу поиска закономерностей) и увеличив обучающую выборку, но для этого необходимо иметь более мощную аппаратно-программную базу и более полную историю прибытия и простоя поездов на станции.

Еще один способ повышения точности прогноза – найти модель, которая бы дополняла нейронечеткую систему и перекрывала ее недостатки. Такими моделями могут стать регрессия или генетические алгоритмы.

Литература

1. Fogler H.R. A pattern recognition model for forecasting // Management science. 1974, No.8. P. 1178 – 1189.

2. Rogelio Oliva, Noel Watson Managing Functional Biases in Organizational Forecasts: A Case Study of Consensus Forecasting in Supply Chain Planning // Mays Business School. 2007, 36p.

3. Чучуева И. Модель прогнозирования временных рядов по выборке максимального подобия. Диссертация [Текст].

4. Пахомова В.Н. Разработка подсистемы оперативного прогнозирования простоев прибывающих поездов на основе системы ANFIS [Текст]/ Пахомова В.Н., Дмитриев С.Ю. // Информационно-управляющие системы на железнодорожном транспорте. – 2013.

5. Marcian N. Cirstea. Neural and Fuzzy Logic Control of Drives and Power Systems [Текст] // Marcian N. Cirstea, Andrei Dinu, Jeen G. Khor, Malcolm McCormick. – Newnes, 2002. – 400 p.

Zhukovytsky I., Pahomova V., Dmitriev S.

The review of methods of prediction which can be used for prediction on railway transport is submitted. For prediction of downtime freight trains for Pyatikhatki's stations it is offered to use an adaptive network of an indistinct conclusion of ANFIS. Implementation of such prediction in a package of the Fuzzy Logic Toolbox expansion of MATLAB system is described.

Keywords: prediction, freight trains, neuro-fuzzy network, training set

Жуковицький І.В., Пахомова В.М., Дмитрієв С.Ю.

Представлено огляд методів прогнозування, які можуть бути використані на залізничному транспорті. Для прогнозування простою вантажних поїздів на станції П'ятихатки пропонується використовувати адаптивну мережу нечіткого виводу ANFIS. Описана реалізація такого прогнозу в пакеті розширення Fuzzy Logic Toolbox системи MATLAB.

Ключові слова: прогнозування, вантажні поїзди, нейронечітка мережа, навчальна вибірка