

УДК 519.2 (34 +54):004.75

Тарек Юсеф Бади Биштави, Г.Н. Жолткевич, Ю.В. Соляник

Харьковский национальный университет им. В.Н. Каразина, Харьков

МНОГОКРИТЕРИАЛЬНОЕ ОБУЧЕНИЕ ПРЕДИКТОРА В ЗАДАЧЕ ПРОГНОЗИРОВАНИЯ ТЕЛЕТРАФИКА

В работе предложено использовать методы многокритериальной оптимизации при построении (обучении) ансамбля предикторов временного ряда. Описан численный эксперимент, в котором показан эффект применения этого подхода на примере построения прогноза характеристик трафика беспроводной сети передачи данных.

Ключевые слова: прогнозирование, многокритериальная оптимизация, машинное обучение, телетрафик, временные ряды, нейронные сети, генетические алгоритмы, компьютерные сети.

Постановка проблемы

Прогнозирование телетрафика играет значительную роль в проектировании и управлении компьютерными сетями. Однако свойства нестационарности, самоподобия и нелинейности его характеристик ограничивают применение эффективных методов стохастического прогнозирования, разработанных для стационарных или специальных случаев нестационарных временных рядов, и делают эту задачу трудновыполнимой [1].

На сегодняшний день известны ряд методов и техник, позволяющие повысить качество прогноза временного ряда за счет использования так называемого «усреднения по ансамблю». Некоторые из них специфичны для модели предиктора:

– метод «усиления» (boosting) [2] позволяет на основе «слабого» предиктора построить коллектив предикторов. Основная идея метода заключается в том, что, при обучении очередного предиктора, основное внимание уделяется тем обучающим примерам, на которых предыдущий предиктор показывает худшие результаты. Затем, полученный в результате голосования результат существенно лучше. Известны работы, в которых этот метод эффективно использовался в прогнозе телетрафика, т.е. в случае нестационарных временных рядов;

– усреднение результатов прогноза группы предикторов. В случае нейронных сетей этот подход естественно использовать, когда веса сети инициализируются случайными числами. Тогда усреднение позволит уменьшить смещение оценки математического ожидания прогнозируемой величины и уменьшить ее дисперсию;

– усреднение результатов работы группы нейронных сетей одинаковой архитектуры, в каждой из которых в скрытом слое содержится избыточное количество узлов (нейронов). Практика показывает, что результаты, полученные с помощью такой техники, близки к результатам, полученным с помощью использования суб-оптимальной (например, полученной перебором) архитектуры нейронной сети.

Очевидно, что основными факторами, влияющими на эффективность использования ансамбля предикторов, являются:

- значение критерия эффективности каждого члена ансамбля,
- взаимная корреляция между остатками (ошибками) каждого из них.

Основная идея

Рассматривается возможность использовать методы многокритериальной оптимизации при построении (обучении) ансамбля предикторов временного ряда. Применение метода «усиления» (boosting) позволяет повысить точность прогноза фактически за счет получения группы предикторов со слабо коррелированными остатками.

Естественно было бы поставить цель явно – при обучении предиктора использовать два критерия: меру точности прогноза и значение коэффициента корреляции между ошибками членов ансамбля.

Обучение нейронной сети в задачах классификации и регрессии сводится к оптимизации некоторой функции – обычно, функции ошибок. Добавление критерия влечет использование методов многокритериальной (многоцелевой) оптимизации [3 – 5].

В этой работе эта идея реализована с помощью генетического алгоритма многокритериальной оптимизации [5], реализованного в пакете Genetic Algorithm and Direct Search Toolbox интегрированной среды вычислений MATLAB. Этот алгоритм реализует один из подходов к многокритериальной оптимизации, использующих понятие так называемой «оптимальности по Парето». Решение Парето-оптимально [4], если значение любого из критериев можно улучшить лишь за счет ухудшения значений остальных критериев.

Далее описан численный эксперимент, в котором показан положительный эффект использования многоцелевого обучения нейронной сети на примере прогнозирования на один шаг вперед характеристик трафика беспроводной сети передачи данных.

Результаты экспериментов

В работе использовались данные характеристик трафика беспроводной сети крупного Интернет-провайдера масштаба города (ЗАО «Виплайн», Воронеж, Россия). К шести точкам доступа провайдера было подключено около 40 беспроводных клиентов. Всего в результате эксперимента было собрано более 12,7 миллионов пакетов. Данные (общий размер Ethernet-кадра за временной интервал в байтах) подробно описаны в [6] и выложены для свободного доступа. В эксперименте (рис. 1) использованы временные ряды, построенные с помощью агрегации (суммирования) исходных значений характеристик трафика по временному интервалу 0,5 с.

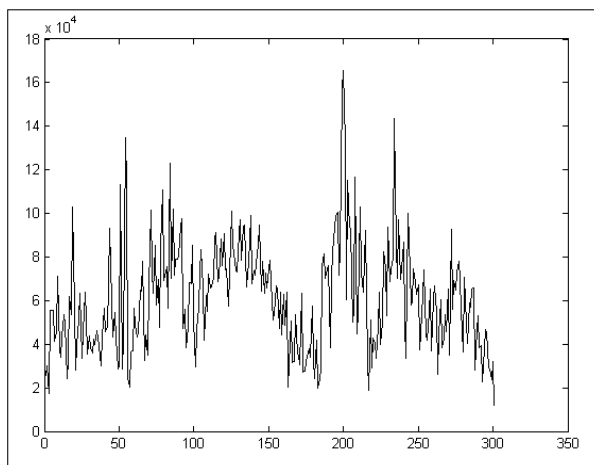


Рис. 1. Фрагмент используемого в экспериментах временного ряда

Для предварительной обработки исходных данных характеристик телеграфика было использовано преобразование, которое позволяет привести ряд к практически стационарному виду, в котором приращения времени сравнимы по абсолютному значению с приращением значений ряда – (логарифм относительного приращения):

$$\ln(x_i/x_{i-1}).$$

Далее данные были «стандартизованы», т.е. приведены к виду,

$$(x_i - m)/s,$$

где x_i – значение временного ряда в i -й точке; m , s – значения оценок математического ожидания и стандартного отклонения.

В качестве модели предиктора временного ряда использована нейронная сеть прямого распространения. На сегодняшний день этот выбор является стандартным и обусловлен тем, что нейронные сети позволяют строить универсальные аппроксиматоры нелинейных функций, которые, в частности, могут симулировать работу большинства известных статистических методов прогнозирования (т.е. заменять их).

В качестве базового предиктора мы использовали нейронную сеть следующей архитектуры:

- 2 слоя (с учетом входного);
- размер 1-го слоя – 3, второго – 1;
- переходная функция второго слоя – линейная (purelin, MATLAB).

Нас интересовала качественная сторона вопроса, поэтому архитектура сети была выбрана простейшая (сеть с меньшим количеством весов требует меньше времени на обучение, а это облегчает воспроизводимость эксперимента). Таким образом, в качестве предиктора использовался линейный фильтр.

Обучение (настройка весов нейронной сети) первого предиктора в ансамбле производилась с помощью генетического алгоритма (использовался пакет Genetic Algorithm and Direct Search Toolbox, интегрированный в систему MATLAB). Кстати говоря, эксперименты, проведенные авторами, показали, что в некоторых случаях (для некоторых видов трафика) при использовании в качестве предиктора многослойного персептрона с нелинейной переходной функцией такой тип обучения дает результаты (точность прогноза на 1 шаг вперед) существенно лучшие, чем, например, алгоритм Levenberg-Marquardt, используемый по умолчанию в пакете Neural Network Toolbox (MATLAB).

В экспериментах 1, 2 ансамбль состоял из десяти предикторов. Для обучения остальных девяти использовался метод многокритериальной оптимизации, реализованный также в пакете Genetic Algorithm and Direct Search Toolbox на основе генетического алгоритма [5].

Целями обучения являлись:

1. Минимизация медианы абсолютных величин ошибок.
2. Минимизация коэффициента линейной корреляции ошибок обучаемого предиктора (на обучающем множестве примеров) с ошибками предыдущего предиктора.

В качестве «предыдущего предиктора» выступали:

1. В одном случае – ансамбль из всех обученных предикторов (*Эксперимент 1* в табл. 1).
2. Во другом случае – последний обученный предиктор (*Эксперимент 2* в табл. 2).

Размер обучающего множества был равен 100 (точек временного ряда). В связи с тем, что одной из важных характеристик предиктора является возможность обобщения, мы испытывали качество предикторов на 100 точках временного ряда, отстоящих с разрывом в 100 точек вперед по отношению к обучающему множеству.

Таблица 1

Эксперимент 1

| | Ансамбль 1 | Ансамбль 2 | Разница в % |
|-------|------------|------------|-------------|
| SER | 8,44 | 8,56 | 1,47 |
| medae | 12401 | 12215 | -1,5 |
| mae | 17535 | 17080 | -2,6 |

Таблица 2

Эксперимент 2

| | Ансамбль 1 | Ансамбль 2 | Разница в % |
|-------|------------|------------|-------------|
| SER | 8,44 | 9,34 | 10,70 |
| medae | 12401 | 12304 | -0,78 |
| mae | 17535 | 16435 | -6,27 |

В качестве мер эффективности использованы:

- среднее абсолютных величин ошибок (mean of absolute values of errors, **mae**);

- медиана абсолютных величин ошибок (**medae**).

- часто используемый показатель точности прогноза «series to error ratio», **SER** (в децибелах – лучшая точность отражается большим значением):

$$SER = 10 \log_{10} \left(\frac{E\{X_t^2\}}{E\{(X_t - \hat{X}_t)^2\}} \right),$$

где $E\{\cdot\}$ – среднее значение; X_t – значения временного ряда; \hat{X}_t – предсказанные значения.

В экспериментах 1 и 2 использовались:

Ансамбль 1 – группа предикторов, при построении которых использовалось одноцелевое обучение.

Ансамбль 2 – группа предикторов, обученных с использованием 2-х описанных выше критериев.

В эксперименте 3 подход к обучению, использовавшийся в эксперименте 2, был испытан для размеров ансамбля 5 и 15.

Таблица 3

Эксперимент 3

| Размер коллектива | 5 (недообучение) | 10 | 15 (переобучение) |
|-------------------|------------------|-------|-------------------|
| SER | -31.20 | 9.34 | -18.90 |
| medae | 42910 | 12304 | 38185 |
| mae | 376608 | 16435 | 145322 |

Как видно из табл. 1 – 3, по всем трем показателям точности использование метода показывает лучшие, чем при обычном усреднении результаты.

Мы испытали в тех же условиях (эксперимент 4) потенциально более мощную нейронную сеть – многослойный перцептрон с нелинейной переходной

функцией (logsig, MATLAB) следующей архитектуры: входной слой – 20; скрытый слой – 5; выходной слой – 1. Обучение сети происходило с помощью генетического алгоритма. В качестве переходной функции использовалась нелинейная функция (logsig, MATLAB). Результаты показаны в табл. 4.

Таблица 4

Эксперимент 4

| | Ансамбль из 10 предикторов | Многослойный перцептрон |
|-------|----------------------------|-------------------------|
| SER | 9,34 | 7,37 |
| medae | 12304 | 15249 |
| mae | 16435 | 20512 |

Выводы

Экспериментальным путем было показано, что использование многокритериальной оптимизации при построении ансамбля предикторов может быть эффективно.

Список литературы

1. Шелухин О.И. Фрактальные процессы в телекоммуникациях: монография / О.И. Шелухин, А.М. Теняшев, А.В. Осин; под ред. О.И. Шелухина. – М.: Радиотехника, 2003. – 480 с.
2. Рассел С. Искусственный интеллект: современный подход / С. Рассел, П. Норвиг. – 2 изд. – М.: Издательский дом «Вильямс», 2007. – 1408 с.
3. Подиновский В.В. Парето-оптимальные решения многокритериальных задач / В.В. Подиновский, В.Д. Ногин. – М.: Наука, 1982. – 256 с.
4. Волгин Л.Н. Принцип согласованного оптимума / Л.Н. Волгин. – М.: Сов. радио, 1977. – 144 с.
5. Kalyanmoy Deb. Multi-Objective Optimization using Evolutionary Algorithms / Deb Kalyanmoy. – Chichester, England: John Wiley & Sons, Ltd., 2001.
6. Петров В.В., Платов В.В. Исследование самоподобной структуры телетрафика беспроводной сети / В.В. Петров, В.В. Платов // Радиотехнические тетради – М.: ОКБ МЭИ, 2004. – № 3. – С. 58-62.

Поступила в редакцию 19.03.2010

Рецензент: д-р техн. наук, проф. И.В. Чумаченко, Национальный аэрокосмический университет им. Н.Е. Жуковского "ХАИ", Харьков.

БАГАТОКРИТЕРІАЛЬНЕ НАВЧАННЯ ПРЕДИКТОРА В ЗАДАЧІ ПРОГНОЗУВАННЯ ТЕЛЕТРАФІКА

Тарек Юсеф Баді Біштаві, Г.М. Жолткевич, Ю.В. Солянік

У роботі запропоновано використовувати методи багатокритеріальної оптимізації при побудові (навчанні) ансамблю предикторів часового ряду. Описано чисельний експеримент, в якому показано ефект застосування цього підходу на прикладі побудови прогнозу характеристик трафіку бездротової мережі передачі даних.

Ключові слова: прогнозування, багатокритеріальна оптимізація, машинне навчання, телетрафік, часові ряди, нейронні мережі, генетичні алгоритми, комп'ютерні мережі.

USING MULTICRITERIA MACHINE LEARNING IN THE TELETRAFFIC FORECASTING PROBLEM

Tariq Usef Badi Bishtavi, G.M. Zholtkevich, U.V. Solynnik

In this paper we proposed to use multi-criteria optimization methods in the construction (training) of the time series predictors ensemble. The numerical experiment shows the effect of this approach on the example of forecasting wireless data network traffic characteristics.

Keywords: prognostication, multicriterion optimization, computer-aided instruction, teletraffic, temporal rows, networks of neurons, genetic algorithms, computer networks.