

ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ МЕТОДИК ПРОГНОЗИРОВАНИЯ ТРАФИКА

Игнатенко Е.Г., Шахов Д. С.

Научный руководитель: к.т.н., доц. Воропаева В.Я.
Донецкий национальный технический университет,
Кафедра автоматизации и телекоммуникаций
ул. Артема 58, г. Донецк, 83001, Украина
Тел.: +380501726584; e-mail: dimshahov@gmail.com

Abstract – in this paper analysis of incoming stream intensity predicting methods is given. Certain advantages and disadvantages of the proposed methods are defined. Most simple and practically applicable prediction models for incoming stream are introduced.

1. Введение

Анализ литературы [1-2] показывает многообразие подходов в прогнозировании поведения информационного потока (экстраполяция, фрактальный анализ, выявление причинной связи и др.). Поэтому вопросы, связанные с исследованием и анализом методов прогнозирования в современных телекоммуникационных сетях, остаются актуальными. Высокая скорость передачи данных и фрактальные свойства трафика диктуют высокие требования к скорости работы алгоритмов прогнозирования. Это поясняет актуальность применения на практике и простых методов прогнозирования.

В докладе приводится анализ и оценки эффективности различных методик прогнозирования на основе регрессионного анализа и скользящего среднего при различной длине выборке предварительных измерений.

2. Основная часть

При выборе методики прогнозирования применяют несколько критериев: горизонт прогноза (краткосрочный, среднесрочный или долгосрочный), доступность статистических данных, структуру данных, точность прогноза. В статье рассмотрены следующие методы прогнозирования:

- 1) линейная регрессия: $Y = a + bx$;
- 2) нелинейная регрессия: $Y = ax^b$;
- 3) множественная регрессия:

$$Y = b_0 + b_1 x_1 + \dots + b_k x_k;$$

- 4) скользящее среднее:

$$Y_{t+1} = \frac{F_t + F_{t-1} + \dots + F_{t-n+1}}{n}.$$

Точность прогноза зависит не только от применяемого метода, но и от используемой при прогнозе длины последовательности статистических данных предыдущих событий. В связи с чем, было необходимо оценить точность всех методов при изменении длины этой последовательности.

Для сравнения применяемых методик, рассмотрено несколько вариантов оценки эффективности алгоритмов прогнозирования, таких как средняя ошибка, среднеквадратическая ошибка, количество ошибок недооценки и переоценки, коэффициенты недооценки и переоценки [3].

Для исследования представленных методов использовалась информация log-файлов, представляющая собой последовательность http-запросов (агрегированная по 1 секунде). Эти данные являются исходными для всех способов прогноза.

В качестве эксперимента в работе рассмотрены 3 сценария, согласно которым прогнозируемая оценка Y_{t+1} определяется на основании некоторого числа предшествующих измерений, т.е. в первом случае число измерений составляет 10 секунд, во втором случае — 5 секунд, третьем — 15 секунд. На рисунке 1 приведены прогнозируемые значения для всех моделей при $n=10$.

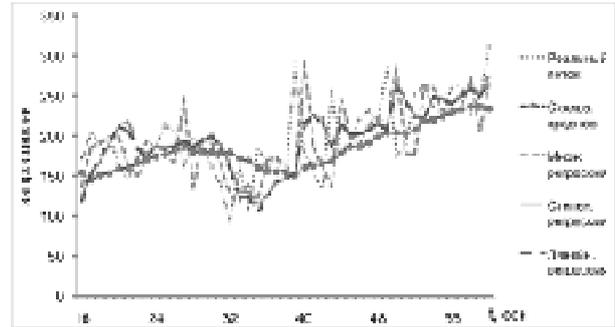


Рисунок 1 — Реальный поток запросов и прогнозируемые значения при $n=10$

Из графика видно, что множественная регрессия наиболее полно повторяет форму кривой интенсивности реального потока, но запаздывает с изменениями при резких изменениях интенсивности запросов. Скользящее среднее сильно сглаживает кривую входящего потока, это связано с количеством отчетов ($n=10$). При уменьшении n скользящее среднее показывает лучшие результаты. Оценки точности нелинейной регрессии улучшаются при увеличении n .

3. Заключение

Анализ приведенных в работе методов прогнозирования показал ряд достоинств: простота реализации, независимость оценок прогнозирования от изменения интенсивности потока. Была оценена точность приведенных методов. Проведенное исследование позволяет определить эффективность применения анализируемых методов в зависимости от длины выборки предварительных измерений и дать рекомендации по выбору метода прогнозирования.

4. Список литературы

1. ГородецкиИ А.Я. Информационные системы. Вероятностные модели и статистические решения: учеб. пособие / А.Я. ГородецкиИ. — СПб: Изд-во СПбГПУ, 2003. — 326 с.
2. ШелухиИ О.И. Фрактальные процессы в телекоммуникациях / О.И. ШелухиИ, А.М. Телякшев, А.В. Осип. — М.: Радиотехника, 2003. — 480 с.
3. Петров В. В. Диссертация на соискание ученой степени кандидата технических наук «Структура телетрафика и алгоритм обеспечения качества обслуживания при влиянии эффекта самоподобия». МосковскиИ ЭнергетическиИ ИнституИ (ТехническиИ университет), 2004г.