

УДК 004.421.2
DOI 10.22213/2410-9304-2018-1-41-44

Р. О. Султанов, кандидат технических наук, доцент
М. О. Еланцев, аспирант
К. А. Орехов, магистрант
ИжГТУ имени М. Т. Калашникова

ПРОГНОЗИРОВАНИЕ ЧИСЛА КЛИЕНТОВ НА ОСНОВЕ КОРРЕЛЯЦИИ СТАТИСТИЧЕСКИХ ДАННЫХ

В статье рассмотрена задача прогнозирования числа клиентов для сервиса обслуживания на примере автомойки. Предложенный алгоритм основывается на поиске корреляции между статистическими данными числа клиентов и количества дневных осадков.

Алгоритм состоит из трех этапов: составление и предварительная обработка сигналов, определение коэффициента корреляции, составление прогноза. Дана оценка эффективности данного алгоритма.

Первый этап включает сбор и предварительную обработку сигналов. Этап основывается на применении фильтров нижних и верхних частот, а также на других многочисленных методах обработки сигналов.

Второй этап основан на определении коэффициента корреляции и его дальнейшего использования в качестве меры отклонения прогноза.

Последний этап состоит из определения пропорций сигналов и применения его при прогнозировании выходного сигнала.

Применение данного алгоритма обработки и прогнозирования сигнала дает возможность работодателям гораздо более эффективно использовать имеющиеся трудовые ресурсы, используя статистические данные числа клиентов и количества осадков.

Ключевые слова: сравнение сигналов, коэффициент корреляции, прогнозирование, обработка сигналов.

Введение

Проблема неэффективного использования трудовых ресурсов растет день ото дня. Работа на предприятиях, предоставляющих различные услуги по обслуживанию, часто имеет сессионный характер. Таким предприятиям необходим прогноз загруженности на несколько дней вперед, чтобы задействовать только необходимое количество персонала и увеличить эффективность использования трудовых ресурсов.

В данной статье рассматривается задача прогнозирования числа клиентов на примере автомойки. Одними из важнейших факторов влияния на число клиентов в данной предметной области являются погодные условия, среди которых большее влияние оказывает количество осадков. Для решения задачи предлагается алгоритм, который на основе данного прогноза входного сигнала (количества осадков) делает прогноз значительного выходного сигнала (числа клиентов), если известна история обоих сигналов.

Алгоритм состоит из трех этапов:

1. Получение и предварительная обработка истории сигналов.
2. Определение коэффициента корреляции сигналов.
3. Составление прогноза выходного сигнала.

Получение и предварительная обработка истории сигналов

Если представить входные статистические данные в виде непрерывных сигналов, как показано на рис. 1, можно четко проследить прямо пропорциональную зависимость между двумя сигналами. Однако как только значение одного сигнала превышает определенный порог (2 мм на графике осадков), дальнейшее увеличение второго сигнала не происходит, что негативно сказывается на коэффициенте корреляции двух сигналов. Такие значения следует исключить, уменьшив их до порогового значения [1].

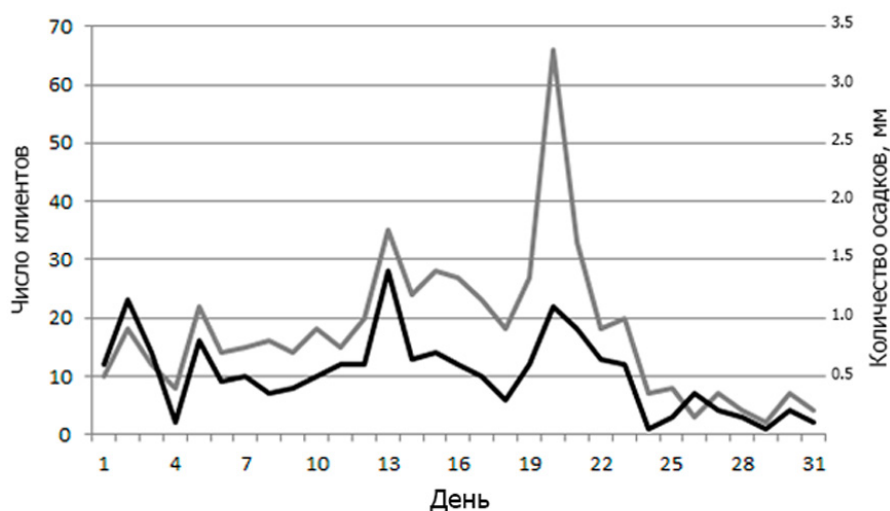


Рис. 1. Графики числа клиентов (черный) и количества осадков (серый)

Определение коэффициента корреляции сигналов

Для того чтобы определить степень влияния входного сигнала на выходной, вычисляется коэффициент корреляции по формуле, разработанной Карлом Пирсоном, Фрэнсисом Эджуортом и Рафаэлем Уэлдоном в 90-х годах XIX века [2]:

$$r_{xy} = \frac{\cos_{xy}}{\delta_x \delta_y} = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum (X - \bar{X})^2 \sum (Y - \bar{Y})^2}}, \quad (1)$$

где $\bar{X} = \frac{1}{n} \sum_{t=1}^n X_t$, $\bar{Y} = \frac{1}{n} \sum_{t=1}^n Y_t$ – среднее значение выборок.

Составление прогноза выходного сигнала

Для составления прогноза выходного сигнала на основе входного необходимо определить, во сколько раз каждое значение первого отличается от второго, иными словами, определить пропорцию p [3]. К примеру, на 1 мм дополнительных осадков приходится 1,2 автомобилей. Значение пропорции p вычисляется по формуле

$$p = \frac{\bar{X}}{\bar{Y}}, \quad (2)$$

где \bar{X}, \bar{Y} – среднее значение выборок.

Далее, имея прогноз входного сигнала, можно вычислить выходной сигнал производством каждого отсчета входного сигнала a на пропорцию p . Коэффициент корреляции

r_{xy} используется в качестве ожидаемого отклонения выходного сигнала. Минимальное прогнозируемое значение выходного сигнала высчитывается по формуле

$$\min = apr_{xy}. \quad (3)$$

А максимальное прогнозируемое значение по формуле

$$\max = \frac{ap}{r_{xy}}, \quad (4)$$

где a – прогноз осадков на выбранный день; p – пропорция; r – коэффициент корреляции.

Чем больше коэффициент корреляции, тем меньше разница между минимальным и максимальным прогнозируемым значением и, следовательно, выше точность прогноза [4, 5].

Вывод

Было проведено исследование работы алгоритма на действующей автомойке в течение шести рабочих дней. На рис. 2 показан график реального числа клиентов (левый столбик) и диапазон прогнозируемых значений (правый столбик) на 6 дней.

Из шести прогнозируемых дней составленный прогноз совпал с днями 2 и 3. Прогноз оставшихся дней был близок к реальным показателям. Чтобы сделать прогноз точнее, в дальнейшем планируется учитывать дополнительные факторы, влияющие на число клиентов.

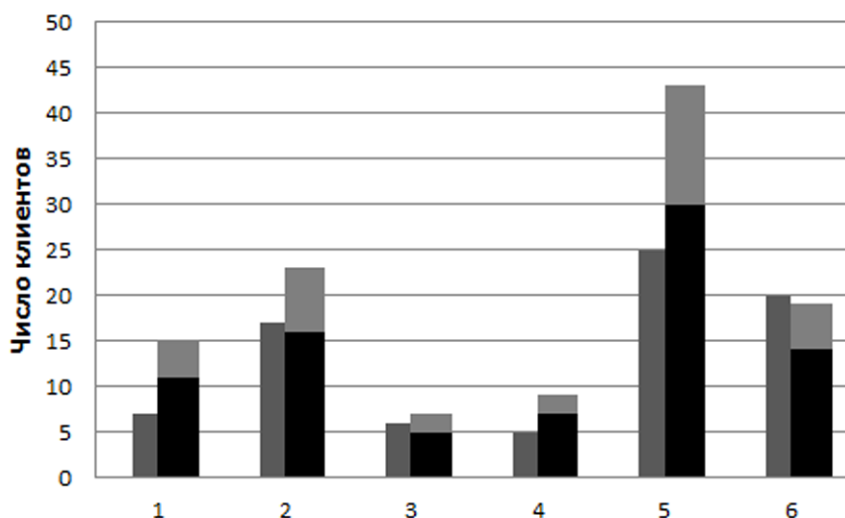


Рис. 2. График реального числа клиентов (левый столбик каждого дня) и диапазон прогнозируемых значений (правый столбик каждого дня)

Библиографические ссылки

1. Айфичер Э. Цифровая обработка сигнала: практический подход. 2-е изд. / пер. с англ. М. : Вильямс, 2004. 992 с
2. Харченко М. А. Корреляционный анализ : учеб. пособие для вузов. Воронежский государственный университет, 2008. С. 13–17.
3. Романенко И. В. Социальное и экономическое прогнозирование. СПб. : Изд-во Михайлова В. А., 2000. 64 с.
4. Статистическое моделирование и прогнозирование / под. ред. А. Г. Гранберга. М. : Финансы и статистика, 1990. 382 с.
5. Халафян А. А. STATISTICA 6. Статистический анализ данных. 2-е изд. М. : Бинوم, 2009.

References

1. Aificher E. (2004). *Cifrovaia obrabotka signala: prakticheskii podhod* [Digital Signal Processing: A Practical Approach], 2nd edition. Moscow: Vil'yams. 992 p. (in Russ.).
2. Harchenko M. A. (2008). *Correliacionnii analiz* [Correlation analysis] Voronezh state University, pp. 13-17 (in Russ.).
3. Romanenko I. V. (2000). *Socialnoe I ekonomicheskoe prognozirovanie* [Social and economic forecasting]. SPb.: Izd-vo Mikhailova V. A. 64 p. (in Russ.).
4. *Statisticheskoe modelirovanie I prognozirovanie* [Statistical modeling and forecasting] (eds. A.G. Granberg). Moscow: Finansy i statistika, 1990. 382 p. (in Russ.).
5. Haliafan A. A. (2009). *Statistica 6. Statisticheskii analiz danih* [STATISTICA 6. Statistical analysis of data. Second edition]. Moscow: Binom (in Russ.).

R. O. Sultanov, PhD in Engineering, Associate Professor, Kalashnikov ISTU
 M. O. Elantsev, Post-graduate, Kalashnikov ISTU
 K. A. Orekhov, Master's Degree Student, Kalashnikov ISTU

Forecasting the Number of Customers on the Basis of Correlation Statistics

The paper considers the problem of forecasting the number of customers to service by the example of a car wash. The proposed algorithm is based on finding correlations between the statistical data on the number of clients and amount of daily precipitation. The algorithm consists of three phases: preparation and preliminary processing of the signals, determination of the coefficient of correlation, and predictive analysis. In conclusion, the estimation of the effectiveness of this algorithm is presented.

The first stage involves gathering and pre-processing. The stage is based on the application of filters of the lower and upper frequencies, as well as on other numerous methods of signal processing.

The second step is based on determining the correlation coefficient and its further use as a measure of forecast deviation.

The last step consists of determining the proportions of the signals and applying it to predict the output signal.

The use of this processing algorithm and the prediction signal gives employers a much more efficient use of available labor resources using statistical data of clients and amount of precipitation

Keywords: comparison of signal, correlation coefficient, forecasting, signal processing.

Получено: 02.03.18