

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИНСТИТУТ ОПТИКИ АТМОСФЕРЫ СО РАН им. В.Е. ЗУЕВА



НОВЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ИССЛЕДОВАНИИ СЛОЖНЫХ СТРУКТУР

**МАТЕРИАЛЫ
ДВЕНАДЦАТОЙ КОНФЕРЕНЦИИ С МЕЖДУНАРОДНЫМ УЧАСТИЕМ
4–8 июня 2018 г.**

*Мероприятие проведено при финансовой поддержке
Российского фонда фундаментальных исследований (проект № 18-07-20033)*

Томск
Издательский Дом Томского государственного университета
2018

Эту дополнительную информацию можно использовать при анализе времени нахождения ВС в состоянии высокой производительности и, не усложняя модели, выработать рекомендации для повышения эффективности работы ВС.

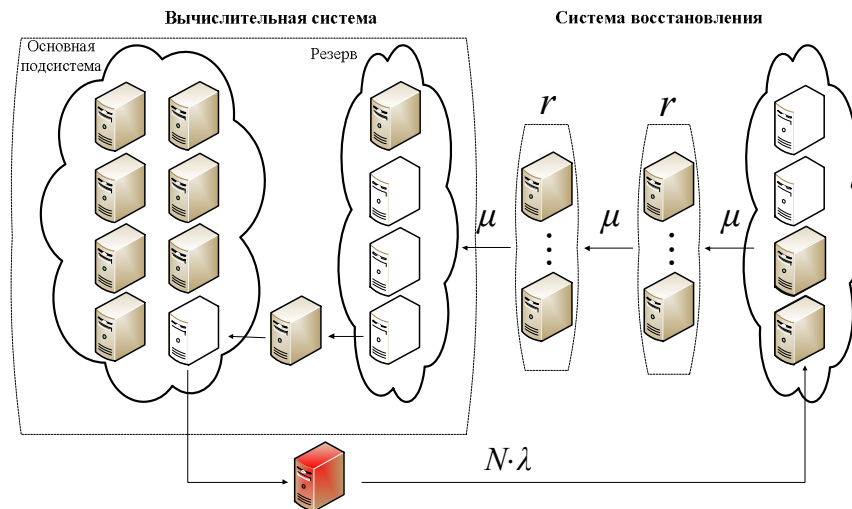


Рис. 1. Модель функционирования ВСс резервом при групповом восстановлении отказавших машин

В работе предлагается аналитическое решение для нахождения функции распределения времени пребывания вычислительной системы в состоянии низкой производительности.

Литература

1. Хорошевский В.Г. Архитектура вычислительных систем. М. : МГТУ им. Баумана, 2008. 520 с.
2. Schroeder B., Gibson G. A large-scale study of failures in high-performance computing systems // Dependable Systems and Networks (DSN2006): proceedings of the International Conference. USA: Philadelphia, PA, 2006. P. 10.
3. Nikolic S. High Performance Computing Directions: The Drive to ExaScale Computing // Труды международной научной конференции "Параллельные вычислительные технологии (ПаВТ'2012)". Новосибирск, 2012. URL: <http://pavt.susu.ru/2012/talks/Nikolic.pdf> (дата обращения: 11.01.2018).

ПРОГНОЗИРОВАНИЕ ТЕНДЕНЦИИ МАКСИМАЛЬНОЙ ЦЕНЫ АКЦИИ

Ю.Г. Дмитриев, Ю.Е. Пальянова

Национальный исследовательский Томский государственный университет, Томск, Россия
pje0360@gmail.com

Пусть имеется N временных периодов с одинаковым количеством n дней. Наблюдается максимальное значение цены акции в каждый из n дней. Максимальные цены между днями будем считать независимыми и одинаково распределенными случайными величинами с функцией распределения $F(x)$. Обозначим максимальную цену акции за n дней в k -м периоде через $X_k^{(n)}$ и рассмотрим последовательность $X_1^{(n)}, \dots, X_N^{(n)}$, отвечающую N периодам. При этом функция распределения $G_{(n)}(x) = P\{X_k^{(n)} < x\} = F^n(x)$. Задача заключается в прогнозировании тенденции (увеличении или уменьшении) максимальной цены акции в $N+1$ периоде по сравнению с предыдущим.

Выберем в качестве $F(x)$ двойной экспоненциальный закон $F(x; \mu, \sigma) = \exp\{-\exp\{-(x-\mu)/\sigma\}\}$, $-\infty < x < \infty$, $-\infty < \mu < \infty$, $\sigma > 0$, предложенный в [1] для описания максимальных значений наблюдаемой величины. Упорядочим наблюдения по возрастанию: $X_{(1)}^{(n)} < X_{(2)}^{(n)} < \dots < X_{(N)}^{(n)}$. Порядковые статистики $X_{(k)}^{(n)}$ можно рассматривать как оценки квантилей закона распределения $G_{(n)}(x)$ уровней $p_k = k/(N+1)$, $k = \overline{1, N}$. На основании этого, следуя [2], можно записать, $F_0^n((X_{(k)}^{(n)} - \mu)/\sigma) = p_k + \varepsilon_{(k)}$, где $\varepsilon_{(k)} = u_{(k)} - p_k$, $u_{(k)}$, k - я порядковая статистика выборки из равномерного в $(0,1)$ закона. Отсюда

$$X_{(k)}^{(n)} \approx \mu + \sigma \cdot Y_{(k)}^{(n)} + \varepsilon_{(k)}, k = \overline{1, N}, \quad (1)$$

где $Y_{(k)}^{(n)} = F_0^{-1}(p_k^{1/n}) = (-\ln(-n^{-1} \ln(p_k)))$, $k = \overline{1, N}$, $F_0(x) = F(x; 0, 1)$ с $M\varepsilon_{(k)} = 0$ и матрицей ковариации $\|M\varepsilon_{(k)}\varepsilon_{(j)}\| = V$ для порядковых статистик из равномерного закона в $(0, 1)$. Модель (1) отражает линейную регрессионную зависимость значений максимальных цен акций от квантилей закона распределения $F_0(x)$. Параметры μ и σ находятся по обобщенному методу наименьших квадратов (ОМНК) [2].

Прогнозирование тенденции максимальной цены акции в $N+1$ периоде предлагается проводить следующим образом [3]. Пусть известны максимальные значения цен акций за N периодов, и вычислены квантили $Y_{(1)}^{(n)} < \dots < Y_{(N+1)}^{(n)}$ за $N+1$ периодов. Поставим в соответствие N упорядоченным значениям цен $X_{(k)}^{(n)}$ N квантилей $Y_{(k)}^{(n)}$ из $N+1$ возможных. Наилучшим будем считать такое соответствие, при котором погрешность $SKO = \sqrt{\sum_{k=1}^N (X_{(k)}^{(n)} - \hat{X}_{(k)}^{(n)})^2}$ (где $\hat{X}_{(k)}^{(n)} \approx \hat{\mu} + \hat{\sigma} \cdot Y_{(k)}^{(n)} + \varepsilon_{(k)}$, $k = \overline{1, N}$) будет наименьшей по всем возможным переборам соответствий с учетом порядка, а также $\hat{\mu}$ и $\hat{\sigma}$, вычисляемых по ОМНК для каждого соответствия. Когда такое соответствие найдено, то становится известен «свободный» квантиль $Y_{(k)}^{(n)}$, которому соответствия не нашлось. Подставляя этот квантиль в оцененное уравнение регрессии при наилучшем соответствии, находим прогнозное значение неизвестной цены акции на $N+1$ период и, соответственно, тенденцию цены.

Метод прогноза был апробирован по реальным данным цен акций ПАО «Газпром».

Литература

1. Смирнов Н.В., Дунин-Барковский И.В. Курс теории вероятностей и математической статистики для технических приложений. М.: Наука, 1969. 512 с.
2. Введение в теорию порядковых статистик: пер. с англ. / под ред. А.Я. Боярского. М.: Статистика, 1970. 413 с.
3. Дмитриев Ю.Г., Устинов Ю.К. Математические модели растущих систем // Вычислительные технологии. 2007. Т. 12, № 1. С. 68–75.

ПОСТРОЕНИЕ РЕГРЕССИОННЫХ МОДЕЛЕЙ НАД ЭМПИРИЧЕСКИМИ РАСПРЕДЕЛЕНИЯМИ

О.А. Попова, Б.С. Добронец, А.А. Жмурова

Институт космических и информационных технологий, Сибирский федеральный университет, Красноярск, Россия
OlgaArc@yandex.ru

Регрессионный анализ данных в настоящее время является областью активных исследований, как в общем теоретическом плане, так применительно к различным областям его практического применения.

Проблема повышения точности результатов численного моделирования, снижения уровня неопределенности с учетом всех особенностей эмпирической информации является актуальной задачей и занимает важное место, в том числе, и для повышения уровня доверия лица принимающего решения к результатам моделирования. Многие исследователи отмечают важность и необходимость анализа эмпирической информации уже на этапе предобработки данных. Обоснованно подобранные модели представления данных на этапе предобработки позволяют определить вид входных переменных и осуществить выбор соответствующих процедур и арифметик для последующего моделирования в соответствии с видом неопределенности и объемом имеющейся информации.

Идея, преобразования эмпирических данных на основе применения математических моделей на этапе представления данных и последующего их использования в виде входных и выходных факторов для моделирования способствовало появлению особого вида переменных. Например, использование гистограммных моделей данных в виде входных переменных для регрессионного моделирования способствовало появлению нового понятия гистограммно-значные переменные, которые представляют собой особый вид переменных, где каждому такому объекту (признаку) соответствует распределение, которое может быть представлено в виде гистограммы. Такие переменные изучаются, например, в символьном анализе [2].

В последнее время наблюдается растущий интерес к моделированию и анализу интервально-значных и гистограммно-значных переменных [2, 3]. Однако анализ публикаций по данной теме исследований показал, что существующие методы и подходы к регрессионному моделированию на гистограммно-значных переменных встречают ряд трудностей [3]. Например, для линейных моделей регрессии для этого типа данных отмечается, что ее параметры не могут быть отрицательными. Для определения параметров этой модели, необходимо решить квадратичную задачу оптимизации, при условии неотрицательности ограничений на