ТЕХНОЛОГІЇ УПРАВЛІННЯ РОЗВИТКОМ

УДК 681.3.06

Л.А. Терейковская

Киевский национальный университет строительства и архитектуры, Киев

ПРИМЕНЕНИЕ ДИСКРЕТНОГО ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ ДЛЯ ОПРЕДЕЛЕНИЯ ЧАСТОТНО-ВРЕМЕННЫХ ПАРАМЕТРОВ ЭКСПЛУАТАЦИОННОЙ НАГРУЗКИ WEB-СЕРВЕРА СИСТЕМЫ ДИСТАНЦИОННОГО ОБРАЗОВАНИЯ

Обоснована необходимость разработки модели расчета частотно-временных параметров эксплуатационной нагрузки Web-сервера системы дистанционного обучения. С использованием теории дискретных вейвлет-преобразований разработано математическое обеспечение такой модели. Указаны пути усовершенствования модели.

Ключові слова: Web-сервер, прогнозирование, вейвлет-анализ, дискретное вейвлет-преобразование

Постановка проблемы

В настоящее время повсеместное внедрение в учебных заведениях систем дистанционного образования приобрело вид ярко выраженной и устойчивой тенденции. В большинстве высших учебных заведений дистанционное образование проводится с использованием таких специализированных платформ как: Lotus LearningSpace, Прометей, WebCT CE, WebCT Vista, Blackboard, ATutor, OLAT и Moodle. В отечественных университетах в основном используется платформа Moodle. Например, данная платформа используется в Киевском политехническом институте, Киевском университете строительства и архитектуры, а также университете «Украина». Не зависимо от типа платформы технической основой ее функционирования является Web-сервер, который обслуживает запросы удаленных клиентов-браузеров. Отметим, что платформа Moodle, в качестве Web-сервера использует Apache. Очевидно, что эффективность функционирования всей платформы дистанционного образования непосредственно зависит от надежности Web-сервера, который как и любое техническое средство имеет свои особенности и ограничения, связанные с его нагрузкой. Например, большинство версий Apache в стандартной конфигурации не в состоянии одновременно обслуживать более 30 клиентов. Таким образом, возникает задача прогнозирования нагрузки на Web-сервер, которая может быть решена за счет разработки соответствующих моделей [3]. В связи с тем, что процесс подключения удаленных клиентов к системе дистанционного образования носит ярко выраженный случайный характер, то, по мнению автора для прогнозирования нагрузки целесообразно использовать статистические модели.

Анализ последних достижений и публикаций

Очевидно, что вид и параметры модели прогноза в значительной мере зависят от тех параметров, которые собственно и определяют нагрузку на Web-сервер. Исходя из типовых задач, выполняемых Web-сервером, а также используемых протоколов передачи данных, оценить нагрузку можно по группам параметров, которые характеризуют использование:

1. Аппаратных ресурсов компьютера-сервера – оперативной памяти, жесткого диска, микропроцессора.
2. Ресурсов операционной системы – длину очереди сетевых соединений, количество открытых файлов, количество запущенных процессов и т.д.
3. Сетевых ресурсов – объем входящего и исходящего трафика, размер сетевых пакетов, количество клиентов в единицу времени, количество «не правильных» сетевых пакетов по каждому из используемых сетевых протоколов и т.д.

Заметим, что перечисленные параметры позволяют непосредственно оценить какую-то одну из физических характеристик технической системы "Web-клиент – Web-сервер" и могут быть зарегистрированы на эксплуатации. Для прогнозирования величин этих параметров во времени в основном используются такие статистические модели прогнозирования данных как линейная (1) и полиномиальная (2), экспоненциальная (3), скользящее среднее (4) и экспоненциальное сглаживание (5).

$$P = a + b \times t; \quad (1)$$

$$P = \sum_{i=0}^{K} \left(a_i \times t^i \right) \ ; \quad (2)$$

$$P = a \times e^{k \times t}; \quad (3)$$

$$P_{t+1} = \frac{P_t + P_{t-1} + \ldots + P_{t-n+1}}{n}; \quad (4)$$

$$S_{t+1} = \alpha \times P_t + \left(1 - \alpha\right) \times S_{t-1}; \quad (5)$$

где P – величина параметра, t – время; a, b – коэффициенты линейной регрессии; K – степень аппроксимирующего полинома; a_i – i-ый коэффициент аппроксимирующего полинома; k – коэффициент; P_t – величина параметра в момент времени t; n – размер скользящего окна (определен во временных отсчетах); S_t – сглаженное значение параметра в момент времени t; α – коэффициент экспоненциального сграживания.

При этом в (3) указывается, что в случае формирования длительного прогноза основной недостаток перечисленных моделей заключается в невозможности учета многопериодического характера типовых зависимостей большинства эксплуатационных параметров Web-сервера от времени. В качестве примера типовых зависимостей параметров от времени приведены графики, показанные на рис. 1 и рис. 2.

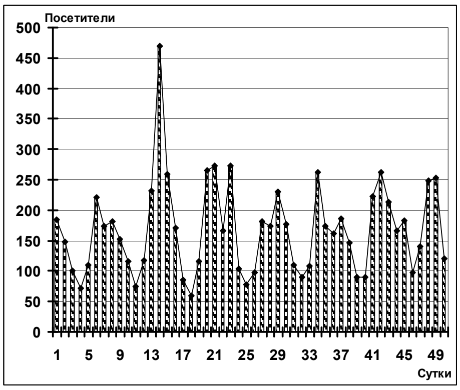

Рис. 1. Зависимость посещаемости Web-сайта от времени

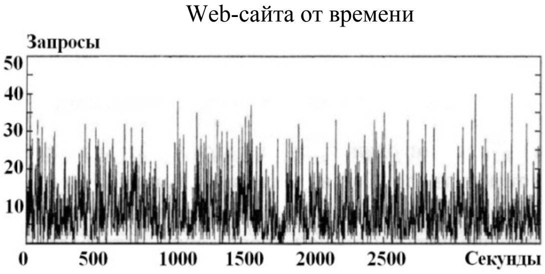

Рис. 2. Зависимость количества запросов к Web-серверу от времени

На рис. 1 показаны график и гистограмма количества посетителей Web-сайта одного из отечественных университетов на протяжении 50 суток при использовании окна регистрации -1 сутки. На рис. 2 показан, приведенный в (3), график зависимости запросов к Web-сайту от времени при использовании окна регистрации 1 секунда.

Даже поверхностный, визуальный анализ рис. 1 и рис.2 указывает на наличие многопериодических зависимостей значений эксплуатационных параметров Web-сервера от времени. Также сравнение графиков посещаемости Web-сайта, построенных с использованием распространенных статистических моделей с реальными данными, подтверждает неадекватность моделей.

Отметим, что на рис. 3 цифрой 1 обозначен график, построенный по реальным данным, а цифрами 2,3,4,5 обозначены графики, построенные с использованием линейной (P=0,3518t+158,71), экспоненциальной (P=139,4exp(0,0035t)), скользящего среднего и полиномиальной моделей (P=0,0051x^3-0,407x^2+9,217x+115,79).

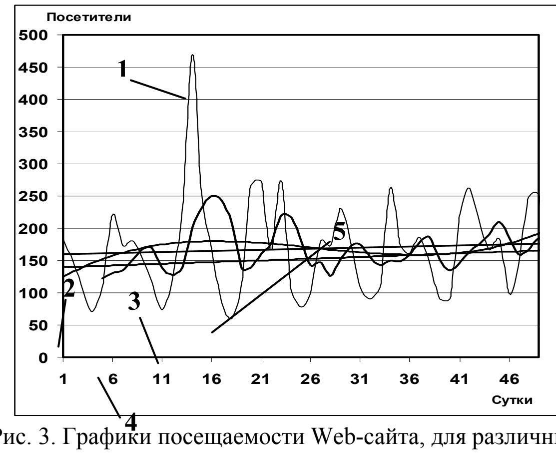

Рис. 3. Графики посещаемости Web-сайта, для различных статистических моделей

Кроме этого, анализ рис. 1 и рис. 2, позволяет с большой долей уверенности предположить, что некоторые периодические составляющие возникают в момент времени $t \neq 0$. Данный факт, а также выводы (2,3) указывают на то, что для прогнозирования эксплуатационной нагрузки Web-сервера необходимо разработать модель расчета частотно-временных характеристик параметров, характеризующих нагрузку.

Одним из наиболее современных методов частотно-временного анализа нестационарных процессов является вейвлет-преобразование. В (3) разработана модель, которая позволяет рассчитать частотно-временные характеристики Internet-серверов. Однако указанная модель базируется на теории непрерывного вейвлет-преобразования, следовательно обладает значительной избыточностью и требует адаптации к дискретным данным, которым соответствует типовая статистика параметров Web-сервера (1, 2).

Возможным путем устранения указанных недостатков может быть использование дискретного вейвлет-преобразования.

Формулировка целей статьи

Используя методы теории дискретных вейвлет-преобразований, разработать математическую модель расчета частотно-временных характеристик эксплуатационных параметров Web-серверов систем дистанционного обучения

Изложение основного материала исследований

В общем случае непрерывное вейвлет-преобразование функции *f(t)* с конечной энергией в пространстве $L^2(R)$ записывается так:

$$W(a,b) = |a|^{-0,5} \int_{-\infty}^{\infty} f(t)\psi^*\left(\frac{t-b}{a}\right)dt, \qquad (6)$$

где W – коэффициент вейвлет-преобразования, ψ – базовый вейвлет (базисная функция); * – процедура комплексного сопряжения; a – масштаб вейвлета; b – сдвиг вейвлета, $a,b \in R$, $a \neq 0$.

При этом, функция ψ должна отвечать таким требованиям: должен быть равным 0 ее нулевой момент, энергия функции должна быть конечной, концентрироваться внутри некоторого финитного интервала и быстро убывать до нуля вне этого интервала (1,2). Для анализа рядов с полиноминальным трендом в базисных вейвлетах должны равняться нулю центральные моменты v-го порядка.

Особенностью дискретного вейвлет-преобразования непрерывной функции является использование дискретных значений масштаба и сдвига вейвлета. Как правило, указанные величины задаются в виде степенных функций вида:

$$a = a_0^{-m}; \qquad (7)$$

$$b = k \times a_0^{-m}; \qquad (8)$$

где m – параметр масштаба; k – параметр сдвига; a_0 – начальный масштаб; m, k, a_0 - целые числа, причем $a_0>1$.

С учетом (7, 8), выражение (6) запишем так

$$W(m,k) = |a_0|^{0,5m} \int_{-\infty}^{\infty} f(t)\psi^*\left(a_0^m \times t - k\right)dt; \qquad (9)$$

Довольно часто a_0 принимают равным 2. Такое дискретное вейвлет-преобразование называют диадным (2). Для диадного вейвлет-преобразования выражения (7, 8, 9) трансформируются так

$$a = 2^{-m}; \qquad (10)$$

$$b = k \times 2^{-m}; \qquad (11)$$

$$W(m,k) = 2^{0,5m} \int_{-\infty}^{\infty} f(t)\psi^*\left(2^m \times t - k\right)dt. \qquad (12)$$

Процедура дискретного вейвлет-преобразования начинается с начального масштаба $a = a_0^{-m}$, которому соответствует уровень минимально допустимого временного разрешения сигнала. Процедура продолжается вместе с дискретным увеличением масштаба, за счет дискретного уменьшения параметра m. Таким образом, анализ начинается с высоких частот и проводится в сторону низких частот. Первое значение масштаба соответствует наиболее сжатому вейвлету. При увеличении величины масштаба вейвлет растягивается. При этом, растяжение вейвлета в a раз по горизонтали приводит к его сужению в a раз по вертикали. В начале анализа вейвлет помещается в начало сигнала (t=0), перемножается с сигналом, интегрируется на интервале своего задания и нормализуется на $|a_0|^{0,5m}$. Результат вычисления *W(a,b)* помещается в точку ($a = a_0^{-m}, b = 0$) масштабно-временного спектра преобразования (2, 4). Сдвиг b может рассматриваться как время с момента t=0, при этом координатная ось b повторяет временную ось функции. Затем вейвлет масштаба a_0^{-m} сдвигается вправо на значение $b = k \times a_0^{-m}$ и процедура повторяется. На частотно-временной плоскости получаем значение, соответствующее $t = b$ и $a = a_0^{-m}$. Процедура повторяется до тех пор, пока вейвлет не достигнет конца сигнала. Для вычисления следующей масштабной строки значение a, дискретно увеличивается на некоторое значение, определяемое параметром $-m$. Тем самым осуществляется дискретизация масштабно-временной плоскости. Максимальное значение масштаба a соответствует продолжительности всего анализируемого ряда данных. Для детализации самых высоких частот сигнала минимальных размер окна вейвлета не должен превышать периода самой высокочастотной гармоники. Если в сигнале присутствуют спектральные компоненты, соответствующие текущему значению a, то интеграл произведения вейвлета с сигналом в интервале, где эта спектральная компонента присутствует, дает относительно большое значение. В противном случае – произведение мало или равно нулю, т.к. среднее значение вейвлетной функции равно нулю. С точки зрения проведения дискретного вейвлет-анализа, статистические данные Web-сервера обладают следующими особенностями:

- ограничены интервалом $t \in [t_{min}, t_{max}]$;
- регистрация данных осуществляется с определенной дискретностью Δt;

- Для диадного преобразования количество точек ряда N должно равняться

$$N = 2^z, \qquad (13)$$

где z – целое число.

С учетом указанных особенностей выражения (9, 12) трансформируются так

$$W(m,k) = a_0^{0,5m} \sum_{i=1}^{N} \left(f(t_i)\psi^*(a_0^m \times t_i - k) \right); \qquad (14)$$

$$W(m,k) = 2^{0,5m} \sum_{i=1}^{N} \left(f(t_i)\psi^*(2^m \times t_i - k) \right); \qquad (15)$$

где N – количество точек ряда, $f(t_i)$ – значение ряда данных в момент времени t_i (i-ый момент регистрации).

Отметим, что в (15) при фиксированном масштабе сдвиг изменяется в интервале от 0 до 2^{z-m}, а максимальный масштаб равен $z-1$.

Результат дискретного вейвлет-преобразования, т.е. матрицу $W(m,k)$, называют вейвлет-спектром.

Обратное дискретное вейвлет-преобразование, которое по своей сути является восстановлением функции по набору вейвлет-коэффициентов определяется выражением:

$$f = \frac{\pi}{ln\, a_0} \sum_{m=0}^{N-1} \sum_{k=0}^{L-1} \psi^* W_{m.k} \ , \qquad (16)$$

где L – количество масштабов.

С целью верификации моделей (7-16) на реальных данных проведен диадный вейвлет-анализ статистики посещаемости Web-сайта, график которой показан на рис. 1. С учетом необходимости выполнения (13), в процессе анализа были задействованы только первые 32 точки ряда. При этом, z=5, а m_{max}=4. Использован базовый вейвлет Добеши. Расчеты велись с использованием пакета Mathcad. В результате анализа получены, представленные на рис.4 графики вейвлет-спектра.

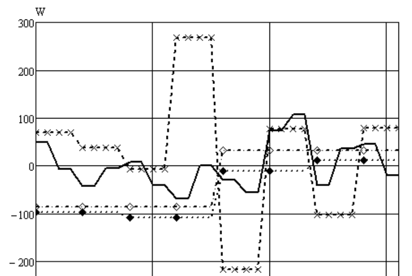

Рис. 4. Вейвлет-спектр посещаемости Web-сервера

Анализ данных рис. 4 указывает на то, что в исследуемом процессе прослеживаются четыре периодические составляющие, которые равняются 2, 4, 8 и 16 суткам. Также можно заметить, что третья составляющая возникает примерно на 14-16 сутки эксплуатации, а остальные не имеют четкой локализации во времени. Полученные результаты свидетельствуют о перспективности расчета частотно-временных характеристик посещаемости Web-сервера с помощью дискретных вейвлетов.

Также проведено обратное дискретное вейвлет-преобразование сигнала по набору вейвлет-коэффициентов. Графики восстановленного и исходного сигналов показаны на рис. 5.

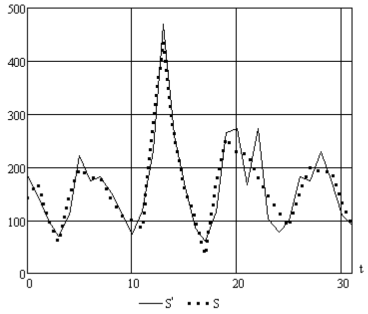

Рис. 5. Графики исходного и восстановленного сигналов

Достаточно высокая схожесть указанных графиков подтверждает возможность применения данного подхода и адекватность разработанных математических моделей. Однако, требуется проведение дополнительных исследований для использования полученных результатов в модели прогноза технического состояния.

Выводы

1. С использованием методов теории дискретных вейвлет-преобразований разработана математическая модель расчета частотно-временных характеристик эксплуатационных параметров Web-серверов систем дистанционного обучения.

2. Перспективы дальнейших исследований в данном направлении заключаются в разработке методики оптимизации вида базисных вейвлетов.

Список литературы

1. Астафьева Н.М. Вейвлет-анализ: Основы теории и примеры применения / Н.М. Астафьева // Успехи физических наук. – 1996. – т.166, № 11 – С. 1145–1170.

2. Дьяконов В.П.Вейвлеты. От теории к практиве / В.П. Дьяконов. – М. :Солон, 2006. – 440 с.

3. Менаске Д. Производительность Web-служб. Анализ, оценка и планирование / Менаске Д., Виргилио А. ; пер. с англ. – СПб. : ДиаСофтЮп", 2003. – 480 с.

4. Михальов О.І. Вейвлет-мультифакторний аналіз складних зображень / О.І. Михальов, Ю.О. Водолазький // Вісник Вінницького політехнічного інституту. – 2009. – № 2 – С. 84–87.

Статья поступила в редколлегию 21.02.2011

Рецензент: д-р техн. наук, проф. С.В. Цюцюра, Киевский национальный университет строительства и архитектуры, Киев