

Терейковська Л. О. (*Держ. унів-т інформаційно-комунікаційних технологій*)

АРХІТЕКТУРА МАРКІВСЬКОЇ МОДЕЛІ ЗМІНИ НАВАНТАЖЕННЯ WEB-СЕРВЕРУ

Терейковська Л. О. Архітектура марківської моделі зміни навантаження Web-серверу. Показано доцільність представлення процесу зміни параметрів навантаження Web-сервера на нестационарних режимах експлуатації у вигляді багатоперіодичного нестационарного процесу. Розроблена архітектура та математичне забезпечення марківської моделі прогнозу навантаження. Для розрахунку ймовірностей переходу та параметрів ланцюга Маркова використано методи теорії вейвлет-перетворень. Визначено умови застосування та вказано перспективні шляхи вдосконалення моделі.

Ключові слова: МАРКІВСЬКА МОДЕЛЬ, ІМОВІРНІСТЬ ПЕРЕХОДУ, WEB-СЕРВЕР, ВЕЙВЛЕТ-ПЕРЕТВОРЕННЯ, ПРОГНОЗ НАВАНТАЖЕННЯ

Терейковская Л. А. Архитектура марковской модели изменения нагрузки Web-сервера. Показана целесообразность представления процесса изменения параметров нагрузки Web-сервера на нестационарных режимах эксплуатации в виде многопериодического нестационарного процесса. Разработана архитектура и математическое обеспечение марковской модели прогнозирования нагрузки. Для расчета параметров цепи Маркова использовались методы теории вейвлет-преобразований. Определены условия использования и указаны перспективные пути усовершенствования модели.

Ключевые слова: МАРКОВСКАЯ МОДЕЛЬ, ВЕРОЯТНОСТЬ ПЕРЕХОДА, WEB-СЕРВЕР, ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЕ, ПРОГНОЗ НАГРУЗКИ

Tereikowska L. O. Architecture to models of the markov of the change the load Web-server. Practicability of the presentation of the process of the change parameter loads Web-server is shown on nonstationary mode to usages in the manner of much periodic nonstationary process is shown. The designed architecture and software to models of the Markov for forecasting of the load is designed. For calculation parameter chain of the Markov methods to theories wavelet-transformations were used. The conditions of the use and is specified perspective ways of the improvement to models are defined.

Key words: MARKOV MODEL, TRANSITION PROBABILITY, WEB-SERVER, WAVELET-TRANSFORMATION, LOAD FORECAST

Практичний досвід та результати [1, 2] вказують на те, що ефективність більшості вітчизняних корпоративних інформаційних систем безпосередньо залежить від якості функціонування Web-серверу, за допомогою якого забезпечуються відповіді на запити віддалених Інтернет-клієнтів. Очевидно, що обчислювальні ресурси Web-серверу повинні відповідати його завантаженості. В протилежному випадку Веб-сервер або не зможе адекватно забезпечувати потреби користувачів інформаційної системи, або резервування надлишкових потужностей призведе до економічних втрат. Крім того висновки і результати, потримані в [1, 2], вказують на невідповідність існуючих інформаційних технологій прогнозування навантаження Web-серверу сучасним вимогам. Разом з тим в [1, 2] доведено, що недостатня точність прогнозування в основному пов'язана з недосконалістю існуючих методів та моделей прогнозу, які базуються на спрощеному математичному апараті і не дозволяють адаптуватись до типових процесів зміни параметрів навантаження.

З точки зору виконання заданих функцій завантаженість Web-серверу залежить від його спроможності обслуговувати HTTP-запити віддалених користувачів корпоративної комп'ютерної мережі. Типовий механізм обслуговування запиту передбачає виконання наступних операцій – отримання HTTP-запиту, встановлення HTTP-з'єднання, пошук цільового ресурсу (файлу), виконання за необхідністю серверних сценаріїв, передача відповіді та розрив з'єднання. В першому наближенні можна вважати, що Web-сервер витрачає однаковий обсяг обчислювальних ресурсів на виконання кожного запиту. Тому навантаження серверу в основному буде залежати від кількості отриманих HTTP-запитів та встановлених HTTP-з'єднань. Однак дані операції можуть бути деталізовані відповідно протоколу TCP/IP, який є основною HTTP. Таким чином, навантаження Web-серверу можливо оцінити за допомогою величин функціональних параметрів котрі відповідають обсягам запитів та з'єднань по протоколам TCP/IP та HTTP. Зазначимо, що практичний досвід та результати [2, 3] вказують на складний вигляд процесу зміни вказаних параметрів. В [1...3] доведено, що для Web-серверу типовими є як стаціонарні, так і нестационарні

режими експлуатації з багатоперіодичним характером зміни величин функціональних параметрів. Нестационарність режимів експлуатації проявляється у не стаціонарності моментів виникнення/зникнення періодичних складових. Для прикладу на рис. 1 показано графік кількості відвідувачів Web-сайту за добу на протязі 50 днів.

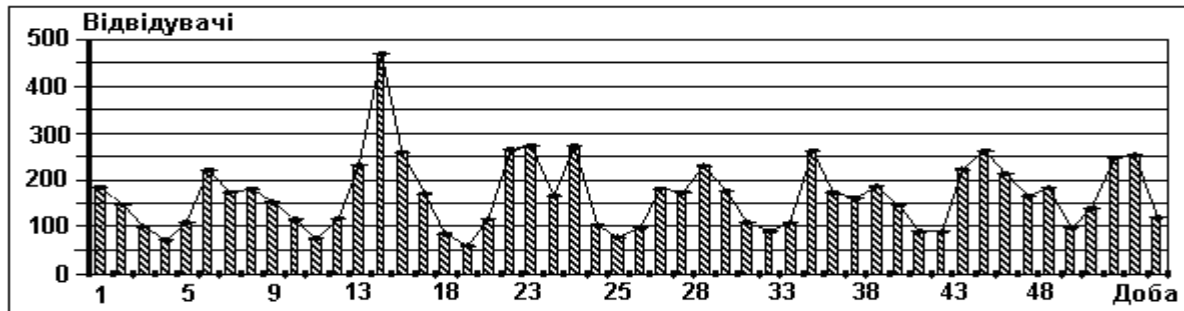


Рис. 1. Зміна кількості відвідувачів Web-сайту

Проаналізуємо типовий процес зміни функціональних параметрів Web-серверу. Відповідно теорії спектрального аналізу даних [4, 5] багатоперіодичний процес можна представити у вигляді сукупності стаціонарних та нестационарних одноперіодичних процесів. Приклади такого представлення показано на рис. 2 і 3. На рис. 2 показано графік двохперіодичного стаціонарного процесу X_3 , який складається з двох процесів X_1 та X_2 , частотні характеристики яких не залежать від часу. На відміну від X_3 двохперіодичний процес X_6 , графік якого показаний на рис. 3, є не стаціонарним, адже до його складу входить одноперіодичний процес X_4 , частота якого залежить від часу.

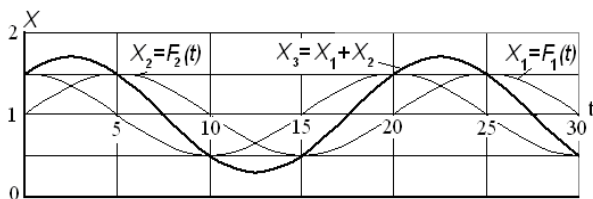


Рис. 2. Двохперіодичний стаціонарний процес

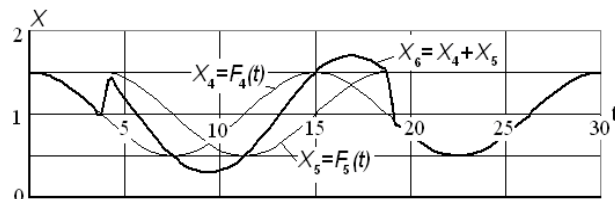


Рис. 3. Двохперіодичний нестационарний процес

Таким чином, в першому наближенні модель прогнозу функціонального параметру Web-серверу в момент t описується виразом:

$$X_{\Sigma}(t) = \sum_{i=1}^K X_i(t), \quad (1)$$

де $X_{\Sigma}(t)$ – величина параметру в момент t ; $X_i(t)$ – значення i -о одноперіодичного процесу; K – кількість одноперіодичних процесів.

Незважаючи на стаціонарність або нестационарність одноперіодичного процесу на кожному з півперіодів його можна представити у вигляді нестационарних процесів 1-го та 2-го роду, графіки яких показані на рис. 4 і 5.

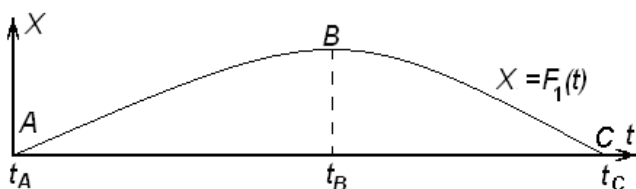


Рис. 4. Нестационарний процес 1-го роду

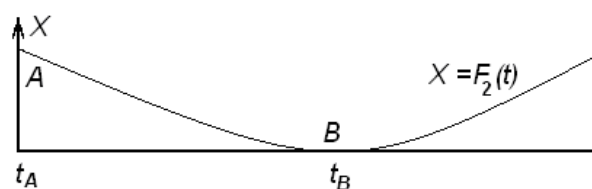


Рис. 5. Нестационарний процес 2-го роду

Вказане припущення дозволяє описати одноперіодичний процес X на протязі всього підслідного інтервалу часу шляхом послідовного використання функцій F_1 та F_2 . Таким чином модель i -го одноперіодичного процесу можна записати у вигляді:

$$X_i = \sum_{j=1}^N (F_{i,1} + F_{i,2}), \quad (2)$$

де N – кількість періодів i -го одноперіодичного процесу; $F_{i,1}$ та $F_{i,2}$ функції, що описуються процеси 1-го та 2-го роду, які відповідають i -му одноперіодичному процесу.

Аналіз нестационарних процесів 1-го та 2-го роду вказує на те, що вони представляють собою композицію монотонно-зростаючого та монотонно-спадаючого процесів, графіки яких показані на рис. 6 та рис. 7:

$$F_1(t) = \begin{cases} F_z(t), t \in [t_A, t_B] \\ F_s(t), t \in [t_B, t_C] \end{cases} \quad (3)$$

$$F_2(t) = \begin{cases} F_s(t), t \in [t_A, t_B] \\ F_z(t), t \in [t_B, t_C] \end{cases} \quad (4)$$

Зазначимо, що для нестационарних процесів 1-го та 2-го роду t_B може не дорівнювати $0,5 \times (t_{A_2} - t_{A_1})$.

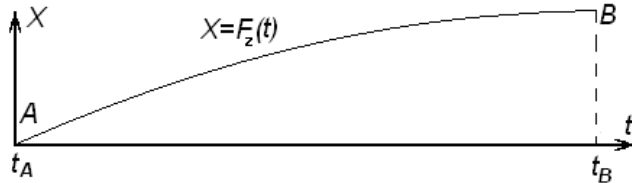


Рис. 6. Монотонний зростаючий процес

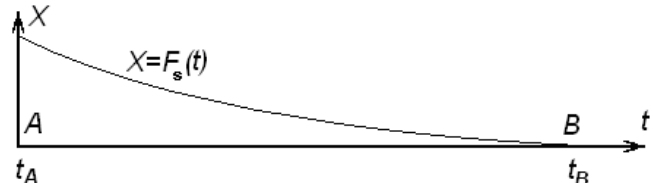


Рис. 7. Монотонний спадаючий процес

Таким чином, аналіз типового процесу зміни функціональних параметрів Web-серверу вказує на те, що модель прогнозу повинна представляти собою ієрархічну структуру, яка складається із окремих блоків пристосованих для моделювання розглянутих процесів. В загальному вигляді математичне забезпечення моделі прогнозу складають вирази (1)...(4). При цьому базовим буде блок моделювання монотонно-зростаючого та монотонно-спадаючого процесів. Зазначимо, що відповідно [2, 6] монотонні зростаючі та спадаючі процеси, графіки яких показані на рис. 6 та 7, адекватно описуються гомогенним ергодичним ланцюгом Маркова з кількістю станів $N=7$. Стани ланцюга Маркова слід трактувати як певні кванти області можливих значень функціональних параметрів. Межі станів ланцюга Маркова можливо розрахувати за допомогою процедури рівномірного квантування. Розрахунок статистичних характеристик процесу у випадку використання марківської моделі здійснюється шляхом рекурсивного вирішення системи рівнянь Колмогорова-Чемпена

$$\begin{cases} P_1(t) = P_1(t-1) - P_1(t-1)p_{1,2} \dots - P_1(t-1)p_{1,7} + P_2(t-1)p_{2,1} \dots + P_7(t-1)p_{7,1} \\ P_i(t) = P_i(t-1) - P_i(t-1)p_{i,1} \dots - P_i(t-1)p_{i,7} + P_1(t-1)p_{1,i} \dots + P_7(t-1)p_{7,i} \\ P_7(t) = P_7(t-1) - P_7(t-1)p_{7,1} \dots - P_7(t-1)p_{7,6} + P_1(t-1)p_{1,7} \dots + P_6(t-1)p_{6,7} \end{cases} \quad (5)$$

де $P_i(t)$ – ймовірність перебування параметру в i -му стану в момент часу t ; $p_{i,j}$ – ймовірність переходу із i -го в j -ий стан за один крок процесу; $N=7$ – кількість станів ланцюга Маркова.

При розрахунках використовується умова нормування
$$\sum_{i=1}^N P_i(t) = 1. \quad (6)$$

Розрахунок перехідних ймовірностей гомогенного ланцюга Маркова реалізується шляхом обробки статистичних даних

$$p(t_k)_{i-1,i} = L(t_k)_i / L(t_k)_r, \quad (7) \quad p_{i-1,i} = \sum_{i=1}^M p(t_k)_{i-1,i} / L, \quad (8)$$

де $p(t_k)_{i-1,i}$ – ймовірність переходу між $(i-1)$ -м та i -м станом на інтервалі t_k ; $L(t_k)_i$ – кількість піддослідних об'єктів, які перейшли із $(i-1)$ -о в i -й стан на інтервалі t_k ; $L(t_k)_r$ – загальна кількість піддослідних об'єктів на момент часу t_k ; $p_{i-1,i}$ – ймовірність переходу між $(i-1)$ -м та i -м станом; L – загальна кількість реєстрацій.

Для моделювання монотонних зростаючих процесів слід застосовувати ланцюг Маркова

$$M_z, \text{ в якому } \begin{cases} p_{i,j} > 0, \exists i > j \\ p_{i,j} = 0, \exists i \leq j \end{cases} \quad (9)$$

Для моделювання монотонних спадаючих процесів слід застосовувати ланцюг Маркова

$$M_s, \text{ в якому } \begin{cases} p_{i,j} = 0, \exists i > j \\ p_{i,j} > 0, \exists i \leq j \end{cases} \quad (10)$$

Враховуючи (3), (4), (9), (10) можливо моделювати нестационарні процеси 1-го та 2-го роду за допомогою послідовного використання двох ланцюгів Маркова M_z та M_s . Ланцюг M_z слід використовувати на інтервалі $[t_A, t_B]$, а M_s – на інтервалі $[t_B, t_C]$. При цьому розподіл ймовірностей M_z в точці B слід використовувати в якості початкового розподілу для M_s . Таким чином марківська модель нестационарного процесу 1-го та 2-го роду описується виразами

$$M_1(t) = \begin{cases} M_z(t), t \in [t_A, t_B] \\ M_s(t), t \in [t_B, t_C] \end{cases} \quad (11) \quad M_2(t) = \begin{cases} M_s(t), t \in [t_A, t_B] \\ M_z(t), t \in [t_B, t_C] \end{cases} \quad (12)$$

Послідовна підстановка виразів (11), (12) в (2) та (1) дозволяє отримати узагальнений математичний опис марківської моделі прогнозу функціональних параметрів Web-серверу. Для практичного використання побудованої моделі необхідно визначити часовий інтервал застосування, розподіл ймовірностей в початковий момент часу ($\|P_0\|$) та ймовірності переходу ($\|p_{i,j}\|$) кожного із ланцюгів Маркова, призначених для моделювання монотонно-зростаючих та монотонно-спадаючих компонентів періодичних складових динаміки параметрів. Визначити часовий інтервал застосування та періодичні складові динаміки пропонується методом вейвлет-перетворень.

Наведемо деякі пояснення. Вейвлет-перетворення являються одним із сучасних розділів теорії спектрального аналізу даних. Важливою перевагою вейвлет-перетворень відносно класичних методів спектрального аналізу є можливість розрахунку частотно-часових характеристик багатоперіодичного процесу. В загальному випадку неперервне вейвлет-перетворення функції $f(t)$ з обмеженою енергією в просторі $L^2(R)$ записується так:

$$W(a,b) = |a|^{-0,5} \int_{-\infty}^{\infty} f(t) \psi^* \left(\frac{t-b}{a} \right) dt, \quad (13)$$

де W – коефіцієнт вейвлет-перетворення; ψ – базовий вейвлет (базисна функція); $*$ – процедура комплексного спряження; a – масштаб вейвлета; b – зсув вейвлета; $a, b \in R, a \neq 0$.

При цьому, функція ψ повинна відповідати таким вимогам: повинен бути рівним 0 її нульовий момент, енергія функції повинна бути обмеженою, концентруватись всередині деякого фінітного інтервалу і швидко зменшуватись до нуля зовні цього інтервалу [4, 5]. Для аналізу рядів з поліноміальним трендом в базисних вейвлетах повинні дорівнювати нулю центральні моменти ν -го порядку.

Неперервні вейвлет-перетворення широко використовуються для аналізу електротехнічних процесів, а для аналізу обмеженого обсягу статистики функціональних параметрів технічних систем застосовуються дискретні вейвлет-перетворення. Їх особливістю являється використання дискретних величин масштабу та зсуву вейвлета. Як правило, вказані величини задаються у вигляді степеневих функцій виду

$$a = a_0^{-m}, \quad (14) \quad b = k \times a_0^{-m}, \quad (15)$$

де m – параметр масштабу; k – параметр зсуву; a_0 – початковий масштаб; m, k, a_0 – цілі числа, причому $a_0 > 1$.

Враховуючи (14), (15), вираз (13) запишемо так:

$$W(m,k) = |a_0|^{0,5m} \int_{-\infty}^{\infty} f(t) \psi^* (a_0^m \times t - k) dt \quad (16)$$

На практиці досить часто a_0 приймають рівним 2. Таке дискретне вейлет-перетворення називають діадним [4, 5].

Для діадного вейвлет-перетворення вирази (14), (15), (16) трансформуються так:

$$a = 2^{-m}, \quad (17) \quad b = k \times 2^{-m}, \quad (18)$$

$$W(m, k) = 2^{0,5m} \int_{-\infty}^{\infty} f(t) \psi^*(2^m \times t - k) dt. \quad (19)$$

Процедура дискретного вейвлет-перетворення починається з початкового масштабу $a = a_0^{-m}$, якому відповідає рівень мінімально допустимого часового дозволу сигналу. Процедура продовжується разом з дискретним збільшенням масштабу за рахунок дискретного зменшення параметру m . Таким чином, аналіз починається з високих частот і проводиться в сторону низьких частот. Перше значення масштабу відповідає найбільш стиснутому вейвлету. При збільшенні величини масштабу вейвлет розтягується. При цьому, розтягнення вейвлету в a раз по горизонталі призводить до його стиснення в a раз по вертикалі. На початку аналізу вейвлет встановлюється в початкову точку сигналу ($t=0$), перемножується з сигналом, інтегрується на заданому інтервалі і нормалізується на $|a_0|^{0,5m}$. Результат обчислення $W(a, b)$ розміщується в точці ($a = a_0^{-m}, b = 0$) масштабно-часового спектру. Зсув b може розглядатись як час, що пройшов з моменту $t=0$, при цьому координатна вісь b повторює часову вісь функції. Після вказаних перетворень вейвлет-масштабу a_0^{-m} зсувається вправо на величину $b = k \times a_0^{-m}$ і процедура повторюється. На частотно-часовій площині отримуємо значення, що відповідає $t = b$ і $a = a_0^{-m}$. Процедура повторюється до тих пір, доки вейвлет не досягне кінця сигналу.

Для обчислення наступного масштабно-часового рядка значення a , дискретно збільшується на величину, що визначається параметром m . Тим самим реалізується дискретизація масштабно-часової площини. Максимальне значення масштабу a відповідає тривалості всього аналізуемого ряду даних. Для деталізації найбільш високих частот сигналу мінімальний розмір вікна вейвлету не повинен перевищувати періоду найбільш високочастотної гармоніки. Якщо в сигналі присутні спектральні компоненти, які відповідають поточному значенню a , то інтеграл добутку вейвлету з сигналом на інтервалі, де ця спектральна компонента присутня, дає відносно велике значення. В протилежному випадку – добуток мало відрізняється від нуля, адже середнє значення вейвлету дорівнює нулю.

З точки зору проведення дискретного вейвлет-аналізу, статистичні дані Web-серверу мають наступні особливості:

- Обмежені інтервалом $t \in [t_{\min}, t_{\max}]$.
- Реєстрація даних реалізується з певною дискретністю Δt .
- Для діадного перетворення кількість точок ряду N повинна дорівнювати

$$N = 2^z, \quad \text{де } z - \text{ціле число.} \quad (20)$$

З врахуванням вказаних особливостей та (20) вирази (16), (19) трансформуються так

$$W(m, k) = a_0^{0,5m} \sum_{i=1}^N (f(t_i) \psi^*(a_0^m \times t_i - k)), \quad (21)$$

$$W(m, k) = 2^{0,5m} \sum_{i=1}^N (f(t_i) \psi^*(2^m \times t_i - k)), \quad (22)$$

де N – кількість точок ряду; $f(t_i)$ – значення ряду даних в момент t_i (i -ий момент реєстрації).

Відзначимо, що в (22) при фіксованому масштабі зсув змінюється в межах від 0 до 2^{z-m} , а максимальний масштаб дорівнює $z - 1$.

Результат дискретного вейвлет-перетворення – матрицю $W(m, k)$, називають вейвлет-спектром.

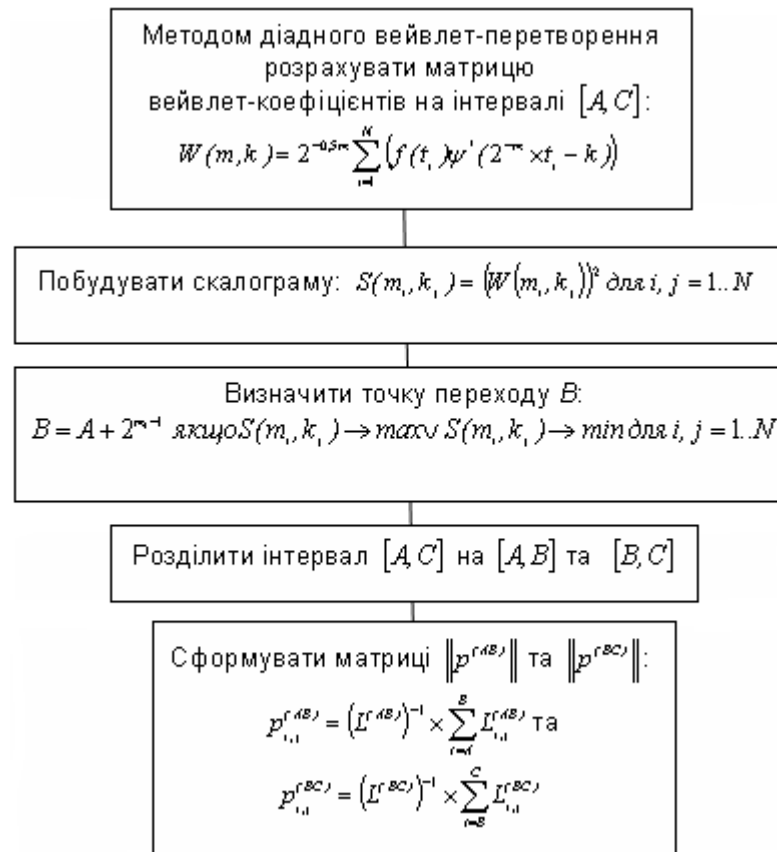
Вейвлет-спектр візуально інтерпретують як поверхню в тривимірному просторі W, a, b . Для виділення найбільш значимих частот розраховують скалограму, яка представляє собою

лінії локальних екстремумів $W(a, b)$ на кожному масштабі a

$$S(a_i, b_i) = |W(a_i, b_i)|^2. \quad (23)$$

Скалограма дозволяє описати розподіл енергії функції вейвлет-спектру по різних масштабах.

Використання діадного вейвлет-перетворення дозволило сформуванню наступної методики розрахунку марківської моделі нестационарних процесів 1-го та 2-го роду:



– За формулами (20), (22) розрахувати матрицю вейвлет-коефіцієнтів.

– За формулою (23) побудувати скалограму.

– Визначити точку переходу процесу, яка відповідає точці максимуму скалограми.

– За формулами (7), (8) розрахувати перехідні ймовірності ланцюгів Маркова призначених для моделювання монотонно-зростаючого та монотонно-спадаючого процесів.

Алгоритм наведеної методики показано на рис. 8.

Висновки

1. Розроблена архітектура та математичне забезпечення марківської моделі прогнозу функціональних параметрів Web-сервера. В моделі передбачено використання для розрахунку ймовірностей переходу та параметрів ланцюга Маркова методів

Рис. 8. Алгоритм розрахунку перехідних ймовірностей

теорії вейвлет-перетворень, що дозволяє адаптувати модель до багато періодичних процесів які відповідають типовій динаміці функціональних параметрів

2. Перспективи подальших досліджень в даному напрямку полягають в розробці загальної методики оптимізації виду базисних вейвлетів.

Література

1. Менаске Д. Производительность Web-служб. Анализ, оценка и планирование / Менаске Д., Виргилио А. ; пер. с англ. – СПб. : ДиаСофтЮп", 2003. – 480 с.
2. Терейковская Л. А. Разработка статистической модели расчета периодических составляющих динамики функциональных параметров Internet-серверов / Л. А Терейковская // Управління розвитком складних систем: зб. наук. праць. – 2010. – Випуск 3. – С. 107-111.
3. Довлад О. А. Дослідження та розробка моделі процесу атаки та трафіку локальної мережі / О. А. Довлад // Захист інформації. – 2009. – № 1. – С. 83-86.
4. Астафьева Н. М. Вейвлет-анализ: Основы теории и примеры применения / Н. М. Астафьева // Успехи физических наук. – 1996. – Т.166, № 11. – С. 1145-1170.
5. Дьяконов В. П. Вейвлеты. От теории к практике / В. П. Дьяконов. – М.: Солон, 2006. – 440 с.
6. Вентцель Е.С. Теория вероятностей / Е. С. Вентцель. – М. : Наука, 1969. – 576 с.