

ОДЕГОВ МИКОЛА

Державний університет інтелектуальних технологій і зв'язку
<https://orcid.org/0000-0001-5526-2487>
onick_64@ukr.net

ГАДЖИЄВ МАТІН

Державний університет інтелектуальних технологій і зв'язку
<http://orcid.org/0000-0001-7280-3863>
gadjievmm@ukr.net

ГЛАЗУНОВА ЛЮДМИЛА

Державний університет інтелектуальних технологій і зв'язку
lvglazun@gmail.com

БУКАТА ЛЮДМИЛА

Державний університет інтелектуальних технологій і зв'язку
ygrikkluda@gmail.com

КОЧЕТКОВА МАРИНА

Державний університет інтелектуальних технологій і зв'язку
jubdv@gmail.com

МЕТОДИКА СТРУКТУРНОЇ ІДЕНТИФІКАЦІЇ МОДЕЛЕЙ НЕСТАЦІОНАРНИХ КВАЗІПЕРІОДИЧНИХ ПРОЦЕСІВ

Задача структурної ідентифікації моделей нестационарних квазіперіодичних процесів вирішуються, як правило, в умовах суттєвої апріорної невизначеності. Заздалегідь невідомо, якими функціями можна моделювати тренд процесу, скільки є приблизно періодичних компонентів, який рівень різного виду шумів спотворює закономірні у середньому процеси. Наша парадигма полягає в тому, що не існує універсального методу для вирішення подібних задач. Тому дослідники мають чітко визначити обмеження застосування методів, алгоритмів і програм обробки даних, що використовуються для вирішення конкретної прикладної задачі. Втім, етапи (кроки) досліджень практично універсальні. Вони є основою сучасних технологій Data Mining, які поєднують неформальний фаховий аналіз даних та формальні операції із застосуванням сучасних методів, алгоритмів та програмного забезпечення.

У даній роботі розглянуто конкретний варіант такого підходу, який включає етапи (кроки) неформального аналізу, формалізації моделі процесу та структурної ідентифікації цієї моделі на прикладі задачі аналізу періодичностей трафіку на великих вузлах обміну (вузлах агрегації трафіку).

Особливостями запропонованої методики є етап (крок) імітаційного моделювання на теоретично подібних до реальних моделей. Мета такого етапу – встановити характеристики ефективності формальних алгоритмів, по можливості обрати в якомусь сенсі найкращий. У формальній частині запропоновано на першому кроці обробки даних виконувати операцію диференціювання вхідного процесу з метою пригнічення регулярного повільного тренду. Шляхом імітаційного моделювання також встановлено, що у спектрах похідних суттєво підкреслюються моди «прихованих» високочастотних періодичностей. Дане положення дозволяє на основі візуального аналізу спектрів визначити частоти зрізу фільтрів нижніх частот, які ефективно пригнічують різного роду шуми, але не відсікають малопомітні регулярні частотні компоненти.

В результаті застосування запропонованої методики очевидні та неочевидні цикли трафіку: річний, піврічний (для північної півкулі), тижневий та добовий. Також виявлені зовсім неочевидні періодичності трафіку: 28 днів та 40 днів. Перший з них корелюється з циклами фаз та орбіти Місяця, до яких пристосувався організм людини на протязі еволюції. Другий поки що можна пояснити лише традиціями та древніми календарями.

Ключові слова: нестационарний квазіперіодичний процес, методи Data Mining, імітаційне моделювання, цифрова фільтрація, спектри похідних процесів, аналіз трафіку

NICK ODEGOV, MATIN HADZHYIEV, LYUDMILA GLAZUNOVA,
LYUDMILA BUKATA, MARINA KOCHETKOVA
State University of Intellectual Technologies and Communication

STRUCTURAL IDENTIFICATION METHODS OF NON-STATIONARY QUASI-PERIODIC PROCESSES MODELS

The task of structural identification of models of non-stationary quasi-periodic processes is solved, as a rule, under conditions of significant a priori uncertainty. It is not known in advance which functions can be used to model the trend of the process, how many periodic components there are, which level of various types of noise distorts regular processes on average. Our paradigm is that there is no universal method for solving such problems. Therefore, researchers should clearly define the limitations of the methods, algorithms and data processing programs used to solve a specific applied problem. However, the stages (steps) of research are practically universal. They are the basis of modern Data Mining technologies, which combine informal professional data analysis and formal operations with the use of modern methods, algorithms and software.

This paper considers a specific version of this approach, which includes stages (steps) of informal analysis, formalization of the process model and structural identification of this model using the example of the task of analyzing traffic periodicities at large exchange nodes (traffic aggregation nodes).

Features of the proposed methodology are the stage (step) of simulation modeling on theoretically similar to real models. The purpose of such a stage is to establish the characteristics of the effectiveness of formal algorithms, if possible to choose the best one in some sense. In the formal part, it is proposed to perform the differentiation operation of the input process at the first step of data processing in order to suppress the regular slow trend. Through simulation modeling, it was also established that the modes of "hidden" high-frequency periodicities are significantly emphasized in the derivative spectra. This provision allows, based on a visual analysis of the spectra, to

determine the cutoff frequencies of low-pass filters, which effectively suppress various types of noise, but do not cut off inconspicuous regular frequency components.

As a result of the application of the proposed methodology, there are obvious and non-obvious traffic cycles: annual, semi-annual (for the northern hemisphere), weekly and daily. Completely non-obvious traffic periodicities were also detected: 28 days and 40 days. The first of them correlates with the cycles of phases and the orbit of the Moon, to which the human body has adapted in the course of evolution. The second can still be explained only by traditions and ancient calendars.

Key words: non-stationary quasi-periodic process, Data Mining methods, simulation modeling, digital filtering, spectra of derivative processes, traffic analysis

Постановка проблеми та аналіз джерел

Як відомо, у строгому розумінні періодичний процес – це функція, яка відповідає залежності:

$$S(t + nT) = S(t), \quad n = 1, 2, \dots, \quad (1)$$

де T – період. Також відомо, що такі функції моделюються рядом Фур'є [1]. Зрозуміло, що у практичних задачах аналізу даних функція виду (1) – лише модель, математична абстракція, яка в окремих випадках може використовуватись для приблизного опису реального процесу (можливо, з достатньою точністю). При цьому теоретична точність обмежується нестабільністю частоти хронометрів, яка, навіть, для прецизійних засобів вимірювань обмежується порядками 10^{-17} [2]. У значній кількості задач умови залежності (1) можуть не виконуватись навіть приблизно. Типові спотворення цієї залежності: наявність аперіодичної складової (тренду) – нестационарність; поступові, або випадкові зміни періодів – аперіодичність; зміни фаз складових – тремтіння фаз, частотні модуляції; наявність різного роду шумів. Тому модель квазіперіодичних нестационарних процесів можна представити у більш загальному вигляді:

$$S(t) = Y(t) + \sum_{k=1}^K H_k(t) + \xi(t), \quad H_k(t) = (h_{k,0} + \Delta h_k(t) + \xi_{H,k}) \cos(\Theta(t) + \xi_{\Theta,k}), \quad (2)$$

де $Y(t)$ – тренд процесу; $H_k(t)$ – квазіперіодичні компоненти (КПК); $\xi(t)$ – адитивна складова шумів процесу; $\Delta h_k(t)$ – регулярна функція зміни амплітуд КПК; $\xi_{H,k}$ – випадковий процес зміни амплітуд КПК (мультиплікаційний шум); $\Theta(t)$ – регулярна функція зміни фаз КПК, яка включає можливі регулярні модуляції частоти і фази; $\xi_{\Theta,k}$ – випадкові викривлення фази (фазовий шум, або «тремтіння» фази). Зауважимо, що функції виду $\Theta(t) + \xi_{\Theta,k}$ включають всі можливі варіанти аперіодичностей.

Модель (2) може бути прийнятною для широкого класу практичних задач: від досліджень явищ мікроміру [3], що стосуються супутникової навігації до дослідження процесів космічного масштабу [4]. У масштабі Земної кулі та окремих регіонів подібні моделі застосовуються для аналізу впливу активності Сонця на формування шарів осадових порід [5], для визначення пікових відхилень атмосферних опадів від звичайних ритмів внаслідок змін клімату [6] та ін. У медичній галузі типовими є задачі автоматизованого аналізу кардіограм [7] та енцефалограм [8]. Подібні моделі також широко застосовуються у технічній сфері, наприклад у теорії складних сигналів [9].

Методи, що застосовуються для ідентифікації моделі (2) залежать від задачі досліджень. Перелічимо найбільш типові з цих задач та визначимо групи відомих методів, що застосовуються для їх вирішення.

1. Задача ідентифікації тренду з метою визначення характерних особливостей процесу та прогнозування (екстраполяції) його розвитку. Методи: ковзного середнього, регресійного аналізу, апроксимації сплайнами [10]. При апріорній невизначеності математичної форми моделі тренду застосовуються методи типу метода групового врахування аргументів [11, 12].

2. Задача спектрального аналізу: по суті, задача розкладання процесу у базисі ортогональних функцій, не обов'язково гармонічних. Методи: класичний спектральний аналіз [1], методи розкладання у базисах функцій Уолша, Хаара, Радемахера та ін. [13].

3. Відокремлення регулярних складових процесу від шумів – задача фільтрації. Методи: на даний час найчастіше використовуються відомі методи цифрової фільтрації [14, 15], а для ряду задач – відносно нові алгоритми вейвлет-аналізу [16].

4. Задача виявлення прихованих періодичностей – типова задача Data Mining [17], яка може вирішуватись, у тому числі, із застосуванням методів кластерного аналізу.

5. Задача виявлення відхилень від нормального розвитку процесу – типова задача діагностики. Методи: з початку треба мати модель «нормального» розвитку процесу, для чого використовуються практично всі перелічені вище методи.

6. Комплексна задача структурної ідентифікації моделі (2) для часткового або повного вирішення всіх названих вище задач. Як і у окремих випадках, що наведені вище, така задача в повному обсязі принципово не може вирішуватись лише за допомогою автоматичних засобів аналізу, наприклад, за допомогою бібліотек аналізу MathLab [14], Phyton [18], Scilab [19] та ін.

Методика, що розглядається у даній роботі полягає у застосуванні основних етапів та методів Data Mining [17, 20], а також відомих методів цифрової обробки даних стосовно задачі ідентифікації моделей виду (2).

Метою роботи є:

- обґрунтування конкретного варіанту послідовності операцій отримання, обробки та аналізу даних на прикладі дослідження прихованих періодичностей;
- визначення та формалізація основних обмежень застосовності пропонованої методики.

Пропонована методика вирішення задачі

В цілому, пропонований у даній роботі варіант методики полягає у послідовному або циклічному виконанні наступних кроків (етапів досліджень).

1. Неформальний експертний аналіз процесу командами фахівців. Мета: формулювання задачі досліджень та виявлення основних характеристик процесу у вигляді вербальної моделі.

2. Формалізація задачі досліджень та основних характеристик моделі.

3. Розробка або вибір алгоритму (методу) структурної ідентифікації моделі для вирішення формалізованої задачі. Розробка програмного забезпечення (ПЗ) для реалізації даного алгоритму або використання стандартних бібліотек ПЗ стосовно розробленого алгоритму.

4. Визначення параметрів точності, екстраполяційної стійкості, фільтраційних можливостей та інших характеристик алгоритму шляхом імітаційного моделювання. При необхідності повернення до кроку 3, або, навіть, до кроків 2 та 1.

5. Застосування розробленого алгоритму та ПЗ для вирішення задачі ідентифікації моделі (2) на реальних даних. Перевірка адекватності отриманих рішень за допомогою зовнішніх критеріїв [11, 12], тобто на додаткових масивах експериментальних даних [17, 20]. При необхідності – повернення до кроку 1.

Дана методика відрізняється від пропонованої раніше [21, 22] включенням додаткових етапів досліджень та врахуванням більшої кількості варіантів випадкових спотворень процесу.

Розглянемо дану методику по крокам на прикладі дослідження процесів росту та коливань трафіку на крупних міжнародних вузлах обміну даними.

Крок 1: Неформальний аналіз

Приклади залежностей трафіку бразильському вузлі (точці) обміну [23] наведені на рис.1 та рис.2 з різними періодами досліджень.

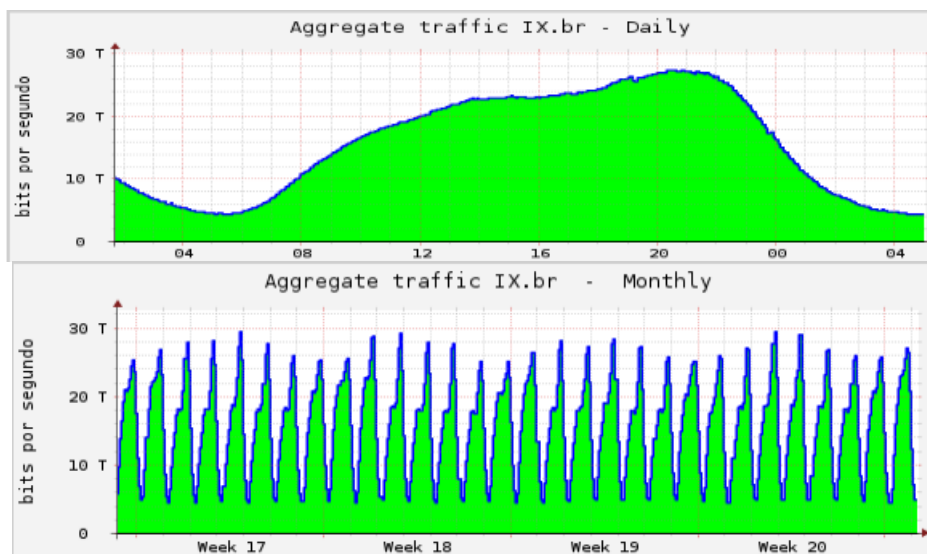


Рис. 1. Добовий та місячний трафік у точці обміну IX.br

Аналіз графіку на рис. 1 дозволяє зробити попередні висновки:

- серед КПК присутні квазіперіоди приблизно у добу та менш помітні періоди приблизно у тиждень;
- на протязі доби трафік змінюється суттєво: від 0,5 Тбіт/с до 25 Тбіт/с (приблизно у 50 разів);
- тижневий цикл менш виразний: пікові значення трафіку за добу коливаються в межах 25–30 Тбіт/с, тобто в межах приблизно 20%.

Аналіз графіків на рис. 2 дозволяє зробити висновки щодо трендів розвитку процесу:

- тренд в усіх випадках є монотонно зростаючою функцією;
- на протязі року тренд зростає приблизно лінійно;
- у більш тривалі періоди лінійний характер тренду вже не є характерним: наприклад, у точці В на рис. 2 спостерігається очевидний перегін тренду: на інтервалі АВ (приблизно 3 роки) трафік зростає приблизно від 1 до 4 Тбіт/с (з темпом приблизно 1 Тбіт/с щорічно); на інтервалі ВС (приблизно 4 роки) трафік зріс приблизно з 4 до 16 Тбіт/с (тобто з темпом приблизно 4 Тбіт/с щорічно).

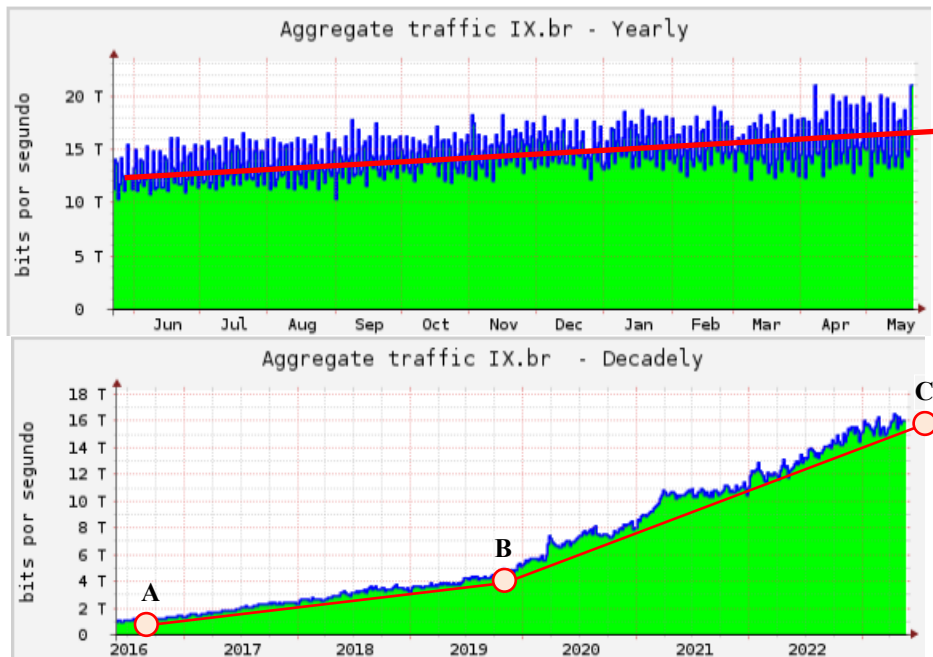


Рис. 2. Трафік у точці обміну IX.br за рік та за 7 років

Зауважимо, що слова типу «приблизно», «на перший погляд» та т. ін. є типовими на даному етапі застосування методики. Надалі ми їх будемо опускати. Також зауважимо, що виявлення особливостей графіків вже на цьому етапі вже дають інформацію для фахівців з прикладних задач. У даному випадку аналіз графіку дозволяє зробити висновки щодо зростання економіки тієї чи іншої держави, щодо розвитку національних мереж передачі даних, або щодо масового втілення новітніх технологій доступу до цих мереж.

Для порівняння наведемо також приклад процесів коливань трафіку у берлінській точці обміну [24]. На рис. 3 показані графіки статистики трафіку за місяць та за рік.

Аналіз графіків на рис. 3 дозволяє зробити висновки, аналогічні до попередніх;

- наявність квазіперіодів тривалістю доба та тиждень; при цьому амплітуда тижневого циклу суттєво менша, ніж добового;
- тренд також має характер зростаючої функції і, можливо, може апроксимуватись поліномом невисокого ступеню.

Втім, помітні і відмінності:

- темпи зростання трафіку значно менше, ніж у бразильській точці (що може свідчити про сталий характер економіки та насичення ринку послуг передачі даних);
- на річному графіку можна помітити цикли, менші, ніж рік, але більші, ніж місяць (сезонні ефекти).

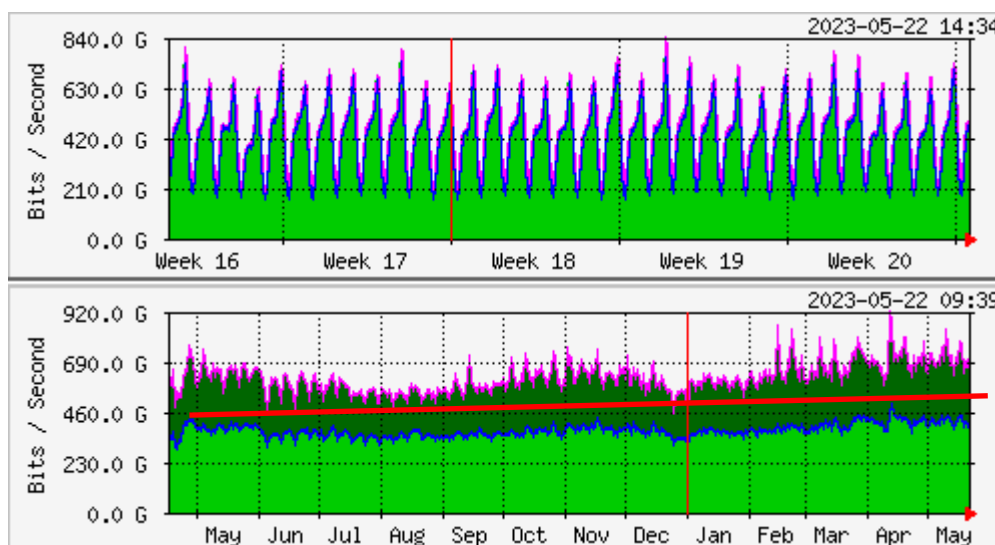


Рис. 3. Трафік у точці обміну BCIX.de за місяць та за рік

Останнє положення характерно для держав, які знаходяться далеко від екватору, як, наприклад Німеччина: соціальні цикли корегуються з природними, а трафік також корегується з цими циклами.

В цілому, аналіз графіків у точках обміну [23, 24], та прикладів для інших точок обміну [21, 22] дозволяє вже при візуальному аналізі зробити висновки про темпи зростання трафіку у середньому (зростання тренду) та визначити найбільш характерні квазіперіоди – добовий та?... А скільки взагалі і які вони хоча б приблизно ці квазіперіоди? На це питання важко відповісти без застосування формальних методів обробки даних, які дозволяють при виконанні певних умов виявити приховані періодичності.

Крок 2: Формалізація задачі та основних характеристик моделі

Для декомпозиції складових моделі (2) на тренд, КПК та адитивний шум треба, щоб у моделі і, відповідно, у даних вимірювань вони якось розрізнялись. Саме від характеристик, які можуть відокремити ці складові і залежить алгоритм, що застосовується для структурної ідентифікації. Для відокремлення тренду та КПК будемо вважати, що тренд є досить повільною функцією часу. Термін «повільний» можна формалізувати по різному. Наприклад, для визначення ступені повільності амплітудних, частотних та фазових коливань пропонувалось використовувати похідні фази [25]. Так, для визначення вузько смугового сигналу пропонувалась оцінка, яку ми уточнимо у термінах та позначеннях задачі, що вирішується:

$$\left| \frac{d^2 \Theta_k(t)}{dt^2} \right| \ll \frac{1}{2\pi} \omega_{0,k}^2, \quad \omega_{0,k} = 2\pi f_{0,k}, \quad f_{0,k} = \frac{1}{T_{0,k}}, \quad T_{0,k} = \frac{1}{p_k} T, \quad (3)$$

де $\omega_{0,k}$ – центральна кутова частота, $f_{0,k}$ – центральна циклічна частота, $T_{0,k}$ – відповідно середній період КПК з номером k ; p_k – середня кількість періодів (не обов'язково ціле число) даної КПК за період спостережень T .

Для ступені повільності тренду використаємо аналогічне визначення: повільним трендом порядку n з параметром ε будемо називати таку функцію $Y(t)$, для якої виконується умова:

$$\frac{\eta(Y^n)}{\eta(Y)} < \varepsilon, \quad Y^n = \frac{d^n Y(t)}{dt^n}, \quad (4)$$

де η – символ оператора норми, яку можна визначити різними способами, як то:

$$\eta_1(Y) = \int_0^T |Y(t)| dt, \quad \eta_2(Y) = \int_0^T Y^2(t) dt, \quad \eta_m(Y) = \max_{t \in [0, T]} |Y(t)|, \quad (5)$$

де η_1 та η_2 – відповідно норми у просторах L_1 та L_2 ; η_m - норма Чебишева.

Всі визначення норм (5) є важливими для задачі, що вирішується. Перші дві з них характеризують відповідно середній розмах значень функцій та їх «енергетичні» характеристики, а норма Чебишева дозволяє порівнювати пікові значення компонентів моделі (2). При необхідності у даній роботі використання тієї чи іншої норми буде уточнюватись.

Якщо теоретично тренд має форму функції, яка допускає розкладання у ряд Тейлора, то параметр ε буде мати сенс максимально можливої норми остатку ряду при отриманні членів включно до ступеню n . Кажучи інакше, тренд можна апроксимувати ступеневим поліномом порядку n з точністю ε незалежно від способу нормування простору.

Операція диференціювання принципово дозволяє відокремити повільний тренд від КПК, які мають приблизно гармонійний характер. При цьому особливо ефективним може бути диференціювання парних порядків: другого, четвертого і т. д. Похідні другого порядку «пригнічують» лінійні складові тренду, похідні 4-го порядку зводять нанівець складові поліномів до 3-го ступеню включно. Навпаки, норми гармонійних складових лише зростають.

Розглянемо простий випадок моделі з трендом у формі поліному другого ступеню та двома гармонійними складовими:

$$S(t) = y_0 + y_1 t + y_2 t^2 + h_{0,1} \cos(\Theta_{0,1} + \omega_{0,1} t) + h_{0,2} \cos(\Theta_{0,2} + \omega_{0,2} t), \quad (6)$$

де модуляційні параметри визначимо так:

$$y_2 = h_{0,1} = h_{0,2} = 1, \quad \omega_{0,1} = 2\pi \frac{1}{T_{0,1}}, \quad T_{0,1} = \frac{1}{4} T, \quad \omega_{0,2} = 2\pi \frac{1}{T_{0,2}}, \quad T_{0,2} = \frac{1}{8} T. \quad (7)$$

Залежності (6) та (7) можуть моделювати процес, який спостерігається 4 роки. При цьому є значні коливання з періодом у рік ($p_1 = 4$ цикли) та півроку ($p_2 = 8$ циклів). Після диференціювання отримаємо значення норм Чебишева похідних всіх складових процесу:

$$\begin{cases} \eta_m(Y^{[2]}) = 2T, & \eta_m(H_1^{[2]}) = 4^2 (2\pi/T)^2, & \eta_m(H_2^{[2]}) = 8^2 (2\pi/T)^2 \\ \eta_m(Y^{[4]}) = 0, & \eta_m(H_1^{[4]}) = 4^4 (2\pi/T)^4, & \eta_m(H_2^{[4]}) = 8^4 (2\pi/T)^4 \end{cases} \quad (8)$$

Залежності (8) показують, що к даному випадку норма другої похідної тренда пропорційна модулю коефіцієнта y_2 , а норма четвертої похідної зведена до нуля (тренд повністю подавлений). На відміну від

цього, норми гармонійних компонентів лише зростають: норми других похідних цих компонентів відрізняються у 4 рази, а норми четвертих похідних – вже у 16 разів. Дане положення тривіальне, але це лише частковий випадок відомих загальних властивостей перетворення Фур'є [1]. Тобто, якщо позначити оператор перетворення Фур'є $F[x]$ якоїсь функції часу $x(t)$, то для похідних цієї функції справедливо: $F[x^{[n]}] = (j\omega)^n F[x]$. Тоді для норми спектру похідних, які зосереджені у локальній полосі $[\omega_k \pm \Delta\omega_k]$ справедливі приблизні оцінки:

$$\eta_1(F[x^{[n]}]) \approx |\omega_k|^n \int_{\omega_k - \Delta\omega_k}^{\omega_k + \Delta\omega_k} |F[x]| d\omega, \quad \eta_2(F[x^{[n]}]) \approx |\omega_k|^n \int_{\omega_k - \Delta\omega_k}^{\omega_k + \Delta\omega_k} (F[x])^2 d\omega, \quad \eta_m(F[x^{[n]}]) \approx |\omega_k|^n \max_{\omega_k \pm \Delta\omega_k} |F[x]|,$$

тобто, **незалежно від способу** визначення норми спектрів КПК будуть пропорційні n -ий ступені частоти для похідних порядку n . Саме дана властивість перетворювання Фур'є і може бути використана для виявлення непомітних на око «прихованих» високочастотних КПК.

Дана властивість похідних високих порядків також дозволяє вирішувати задачу фільтрації шумів. Як відомо, ефективним методом відокремлення «корисного» сигналу від шуму є застосування фільтрів нижніх частот (ФНЧ) [14-16]. Незалежно від функції ФНЧ важливо встановити значення частоти зрізу $\pm \omega_H$ і відповідної ефективної ширини пропускання цього фільтру. Зрозуміло, що у випадку, коли середня частота теоретично найвищої КПК $\omega_{0,K} < \omega_H$, то ФНЧ видалить інформаційні складові з цією частотою:

так би мовити, з водою виплесне ще й дитину. Навпаки, якщо $\omega_{0,K} \gg \omega_H$, то значна доля енергії шумів зостанеться після фільтрації і буде заважати подальшому вирішенню комплексної задачі структурної ідентифікації. Тому візуальне дослідження спектрів **похідних** (кажучи дуже «науково»), метод Visual Mining [17, 20]) може допомогти у вирішенні задачі встановлення оптимальних значень частоти зрізу ФНЧ.

Яка може бути максимальна частота, що принципово можна надійно виділити? Теорема Котельникова-Найквіста стверджує, що сигнал у полосі частот шириною Δf можна однозначно відтворити відліками, які відстоять один від одного на $1/2\Delta f$. У реальних застосуваннях частота відліків береться більшою. Наприклад, модуляції голосу людини мають межу верхньої границі частот 3,5-4 КГц. За згаданою теоремою достатньо відліків з частотою 7-8 Кбіт/с. Втім, у цифровий телефонії обрано стандарт 64 Кбіт/с, що у 8 разів перевищує оцінку Найквіста-Котельникова. З урахуванням можливих шумів різної природи дане значення: 8 відліків на найкоротший період коливань слід прийняти як найменше допустиме значення.

Також для відокремлення залишків енергії тренду після диференціювання корисним може бути фільтрація за допомогою фільтрів вищих частот (ФВЧ) з частотою зрізу $\pm \omega_H$. Якщо похідну тренду певного порядку можна представити приблизно як константу (приклад другої похідної моделі (6)) на інтервалі $[0, T]$, то її амплітудний спектр (надалі – спектр) буде мати вигляд функції $\sin(\omega T/2)/(\omega T/2)$. Найближчі до центру (відповідно, до максимуму) нулі цієї функції будуть у точках $\pm 2\pi/T$ та $\pm 4\pi/T$. При цьому найбільш значна доля енергії похідної тренду буде зосереджена саме на інтервалі $\pm 4\pi/T$, що відповідає двом циклам за час спостережень T . Таким чином, найнижча частота КПК, яку можна відокремити від тренду дорівнює 3-м циклам ($\pm 6\pi/T$).

Таким чином, встановлені основні формальні обмеження застосування пропонуваніх алгоритмів. Вкажемо також обмеження, які важко формалізувати, але легко визначити на інтуїтивному рівні:

- КПК повинні бути відокремлені одна від іншої, хоча б в частині смуги основної енергії;
- для візуального аналізу спектри КПК мають бути приблизно унімодальними;
- тренд має бути повільною функцією порядку n з дуже малим для даної задачі параметром ε ;
- шуми не повинні мати регулярних складових (трендів, помітних кореляцій та ін.).

Втім, природа не повинна підкорятись нашим про неї уявленням. Тому немає іншого способу перевірки адекватності моделей, аніж перевірка їх на додаткових масивах експериментальних даних.

Крок 3: Розробка алгоритму

Операції алгоритму логічно слідує з формального аналізу на кроці 2 та в значній частині включають операції, що було раніше використано [21, 22]. У даній статті ми пропонуємо удосконалений алгоритм.

1. За даними вимірювань отримується масив значень $S(t)$, імовірно у формі залежності (2).
2. Методами чисельного диференціювання знаходиться похідна $S^{[2]}(t)$, або похідні більш високих порядків.
3. Виконується операція перетворення Фур'є, зазвичай алгоритмами швидкого перетворення (ШПФ) $F[S^{[2]}]$. В отриманому спектрі за допомогою візуального аналізу або методами кластерного аналізу Data Mining знаходяться моди КПК з мінімальним квазіперіодом $T_{0,K}$, який включає не менше як 8

відліків. Встановлюється середня частота даної КПК: $\omega_{0,K} = 2\pi / T_{0,K}$ та приблизний інтервал частот, у якому зосереджена більша частина енергії даної КПК: $\omega_{0,K} \pm \Delta\omega_K$. Як правило, для вузько смугових КПК можна обмежитись значеннями $\Delta\omega_K = 0,1\omega_{0,K} \div 0,2\omega_{0,K}$. Встановлюється частота зрізу ФНЧ: $\omega_H \geq \omega_{0,K} + \Delta\omega_K$.

4. Виконується операція ШПФ над *вхідним масивом* значень $F[S]$ та над результатом виконується операція фільтрації ФНЧ, після чого отримується спектр з подавленими вищими частотами: $\Phi_{\text{ФНЧ}}(F[S]) = F^*[S]$. Зворотнім ШПС отримується значення вхідної функції з частково подавленими шумами: $S^*(t) = F^{-1}\{F^*[S]\}$.

5. Шляхом диференціювання отриманої функції додатково подавляються повільні компоненти (очищення від тренду), в результаті чого отримується приблизна оцінка КПК: $H^*(t) = \sum_{k=1}^K H_k^*(t)$.

6. Заключна операція – приблизна оцінка тренду: $Y^*(t) = S^*(t) - H^*(t)$.

Крок 4: Імітаційне моделювання

Ефективність метода імітаційного моделювання полягає в тому, що поряд з еталонною (ідеальною) моделлю процесу є також спотворена шумами, викривлена модель. Тому можна вирішувати зворотню задачу аналізу: застосувавши пропонований алгоритм ідентифікації до викривленої моделі, визначити параметри відхилень результату від еталонної моделі.

Для моделювання на основі залежності (2) визначимо загальний вид моделі, передбачаючи, що «вимірювання» виконувались на протязі 3-х років з кроком у добу (приблизно 1095 відліків). При цьому в ідеалі спостерігались річний (3 цикли), піврічний (6 циклів) та тижневий (приблизно 160 циклів) періоди.

Позначимо: $r(t)$ – значення датчика псевдовипадкових чисел, рівномірно розподілених на інтервалі ± 1 ; $tr(t)$ – значення датчика псевдовипадкових чисел з трикутним розподіленням також на інтервалі ± 1 . Визначимо параметричний вид моделі:

$$\begin{cases} Y(t) = V_0(1+q)^{t/T} \\ H(t) = \sum_{k=1}^3 (h_k + \Delta h_k t/T) \cos(\Theta_{0,k} + 2\pi\Delta\Theta_k r(t) + 2\pi t T/T_k) \\ \xi(t) = A_\xi tr(t) \end{cases} \quad (9)$$

Значення параметрів еталонної та спотвореної моделі за формулами (9) наведені у табл. 1.

Таблиця 1

Значення параметрів моделей

Параметр	Значення для еталонної моделі	Значення для спотвореної моделі
T	1095	1095
q	1	1
h_1, h_2, h_3	0,25; 0,125; 0,01	0,25; 0,125; 0,01
$\Delta h_1, \Delta h_2, \Delta h_3$	0,15; 0,25; 0	0,15; 0,25; 0
$\Theta_1, \Theta_2, \Theta_3$	0; 0,5 π ; 0	0; 0,5 π ; 0
$\Delta\Theta_1, \Delta\Theta_2, \Delta\Theta_3$	0; 0; 0	0,125; 0; 0,01
T_1, T_2, T_3	365; 182,5; 7,0	365; 182,5; 7,0
A_ξ	0	0,15

Із табл. 1 видно, що спотворення еталонної моделі включають фазові шуми та адитивний шум. При цьому розмах коливань адитивного шуму у 15 разів перевищує амплітуду тижневої КПК. Також моделі включають амплітудні модуляції КПК, при цьому суттєві: на протязі інтервалу досліджень амплітуда річного циклу збільшується від 0,25 до 0,40, а піврічного від 0,125 до 0,375 – практично, втричі.

Графіки складових еталонної моделі та сумарний графік спотвореної моделі показані на рис. 4.

У варіанті масштабу на рис. 4 досить легко встановлюється наявність річного та піврічного циклу. Тижневий цикл візуально не помітний на обох графіках. Якщо відобразити фрагмент процесу в іншому масштабі, наприклад, за перші 2 місяці «спостережень» (рис. 5), то на графіку КПК еталонної моделі тижневий цикл добре виражений візуально, а ось на графіку спотвореної моделі даний цикл зовсім не можна побачити на око.

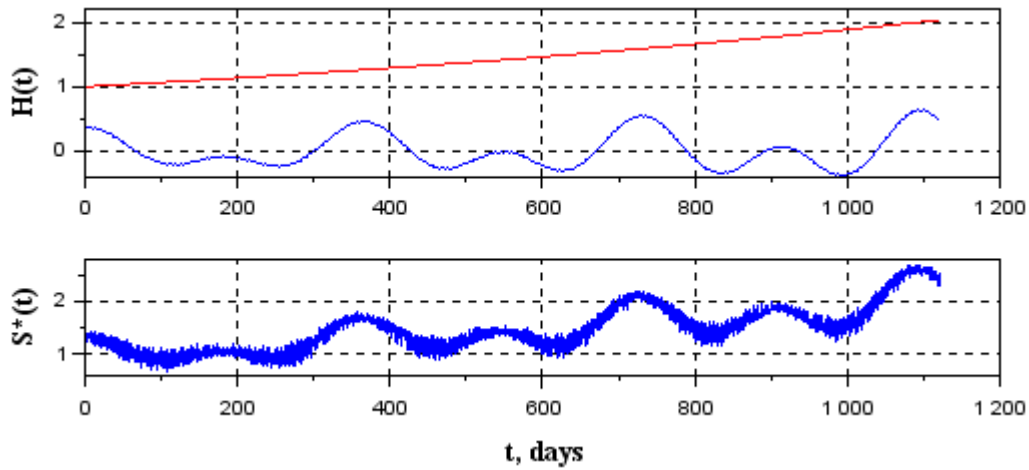


Рис. 4. Тренд та КПК еталонної моделі; сумарний графік спотвореної моделі

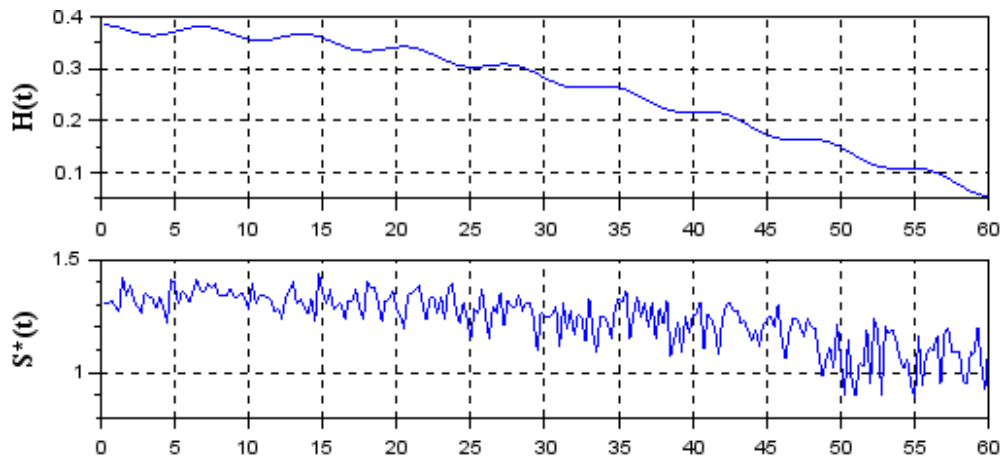


Рис. 5. КПК еталонної моделі та сумарний графік спотвореної моделі за 2 місяці

Таким чином, у даному випадку виявлення тижневої періодичності є типовою задачею Data Mining – задачею виявлення «прихованих» періодичностей. Саме цю задачу і будемо вирішувати за допомогою запропонованого алгоритму.

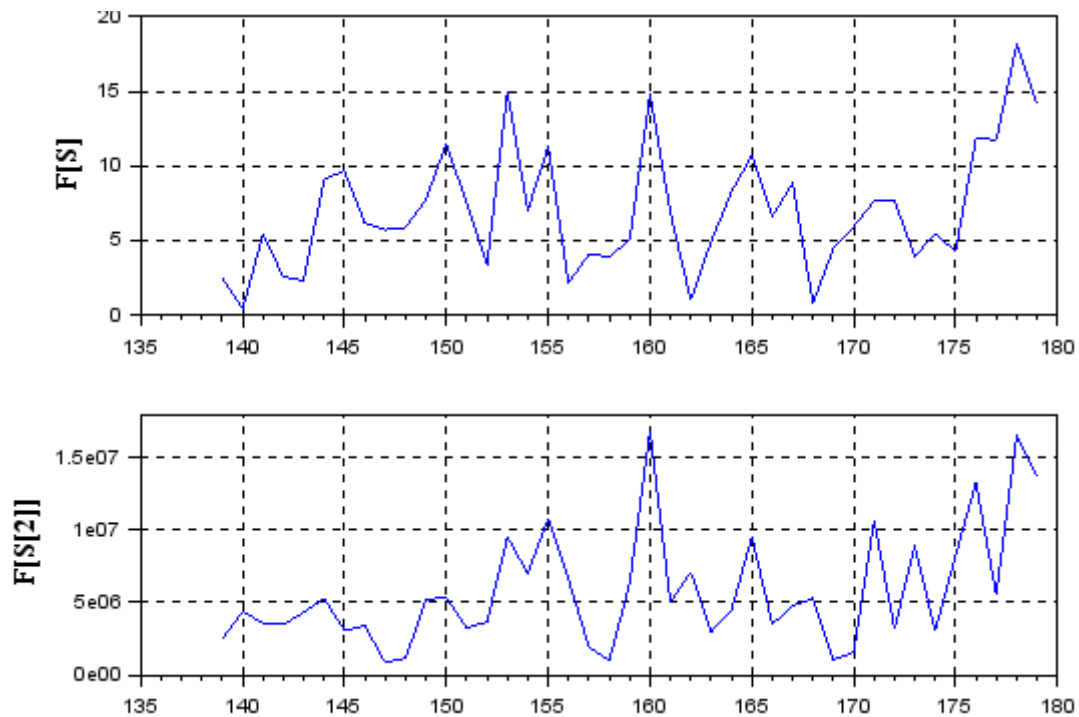


Рис. 6. Спектри спотвореної моделі та її другої похідної в області тижневого циклу

Графіки фрагменту спектру спотвореної моделі $S(t)$ та спектру другої похідної $S^{[2]}(t)$ показані на рис. 6. На цих графіках частота дана у циклах за період досліджень: 160 циклів приблизно відповідає тижневному циклу, який і треба виявити.

Як видно з рис. 6, пік на частоті 160 для $F[S^{[2]}]$, превалює над сусідніми значеннями, які обумовлені спектром шумів. Поряд з цим, у спектрі $F[S]$ є моди, які перевищують значення піку на частоті 160. Втім, це лише одна реалізація випадкового процесу. Для статистичного аналізу треба виконати багаторазове моделювання. Так, на рис. 7 показано результати 20-кратного вирішення тієї ж задачі.

Візуальний аналіз графіків на рис. 7 вже показує, що у середньому пік тижневого циклу більш виражений у спектрі другої похідної. Для статистичного аналізу модальності у даному випадку можна використати показник відносних середніх значень. Для цього відокремимо цикли поблизу тижневої моди 160 ± 2 та цикли більш широкої полоси 160 ± 20 . Позначимо: F_n – амплітуди спектральних складових для відповідних частот. Тоді показник модальності буде визначатись формулою:

$$\eta = \left(\frac{1}{4} \sum_{n=160-2}^{160+2} F_n \right) / \left(\frac{1}{40} \sum_{n=160-20}^{160+20} F_n \right). \quad (10)$$

Значення показника (10) для різних варіантів співвідношення амплітуди тижневого циклу та розмаху значень адитивного шуму дано у табл. 2.

Таблиця 2

Значення показника модальності

Співвідношення A_ξ / h_3	Показник для $F[S]$	Показник для $F[S^{[2]}]$
10	2,46	5,75
15	2,02	4,23
20	1,78	3,42

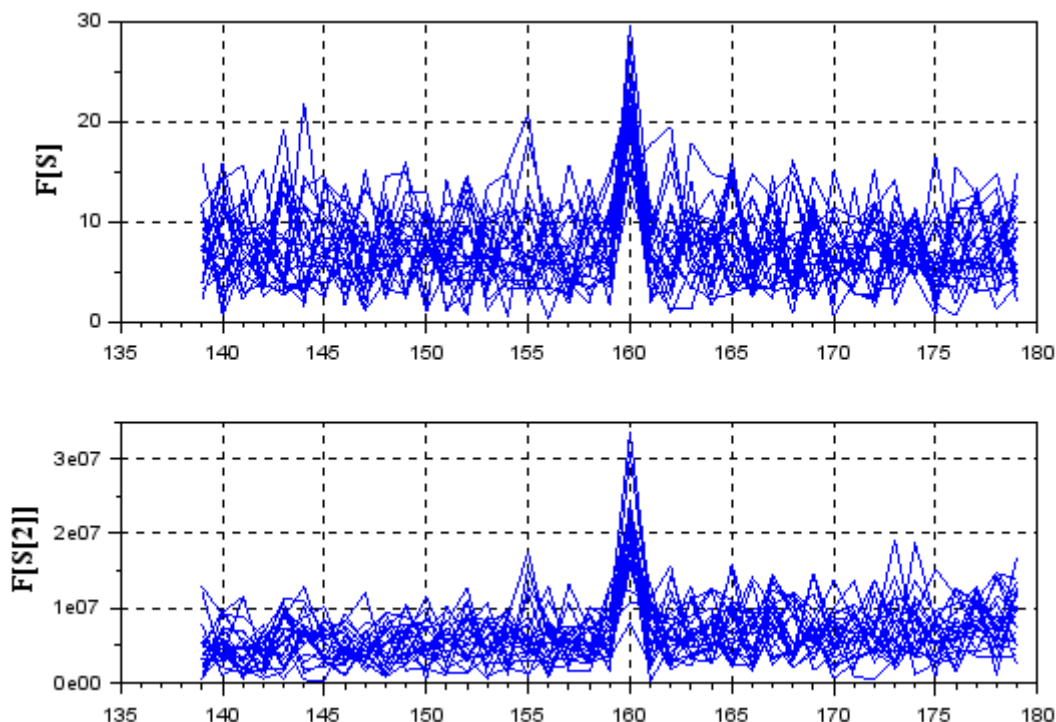


Рис. 7. Спектри спотвореної моделі та її другої похідної в області тижневого циклу

Аналіз даних у табл. 2 дозволяє зробити важливий висновок: показник модальності для спектру другої похідної приблизно вдвічі більший, ніж для спектру самої моделі. Таким чином, операція диференціювання дозволяє підвищити роздільну здатність методів спектрального аналізу у даному випадку.

Після встановлення найвищої частоти «корисних» та значущих КПК виконується крок фільтрації ФНЧ, що показано на рис. 8.

У випадку значень, що відповідають табл. 1, дисперсія адитивного шуму після фільтрації зменшена приблизно у 30 разів. При цьому ефективно визначено моду тижневого циклу. Коментарі зайві.

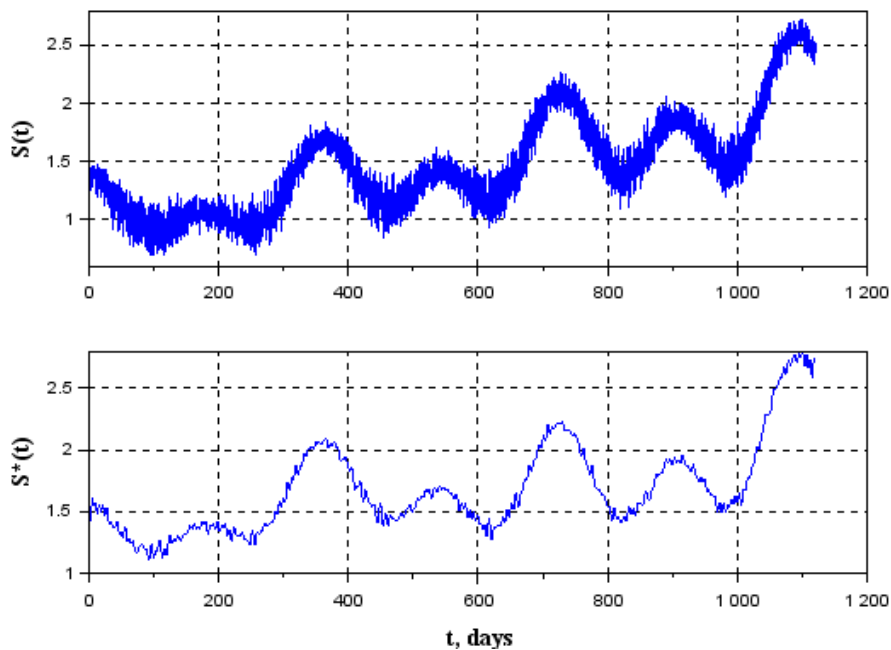


Рис. 8. Спотворена модель до і після фільтрації ФНЧ

Крок 5: Застосування розробленого алгоритму

Дана робота, по суті, є продовженням досліджень, виконаних раніше [21, 22] з урахуванням особливостей використаної методики, а саме: визначення полос фільтрації за допомогою подвійного диференціювання процесів та обов'язкового використання методів імітаційного моделювання на кроці дослідження ефективності алгоритмів вирішення прикладної задачі. Тому опустимо ряд проміжних кроків даної методики і розглянемо лише результати дослідження «прихованих» квазіперіодичних нестационарних процесів на прикладі аналізу трафіку на великих вузлах обміну [23, 24].

Результати спектрального аналізу КПК реальних процесів показані на рис. 9. Для отримання цих результатів виконано попередні операції за кроками методики 1-3 та КПК відокремлені від тренду після фільтрації шумів за допомогою ФНЧ з частотою зрізу, яка на 10% перевищує частоту тижневого циклу. За допомогою ФВЧ додатково подавлені рештки похідних «повільного» тренду порядку більше 2, а також очевидний річний цикл. Графіки показані для досліджень на протязі 3-х років. По вісі ординат показані значення відносно максимуму функцій (умовний масштаб), а по вісі абсцис – частота у циклах на рік.

Порівняльний аналіз графіків на рис. 9 дозволяє зробити певні висновки:

- тижневий цикл (приблизно 52 цикли у рік) помітний на обох графіках;
- піврічний цикл (2 цикли у рік) помітний лише для трафіку у Берліні та на око практично не виражений у бразильській точці обміну;

Втім, особливо цікавими є близькі періодичності у 9 циклів та 13 циклів на рік. Відповідно – приблизно 40 днів та 28 днів. Умовно – місячні цикли. Цікаво, що цикл тривалістю у календарний місяць (30-31 день, або приблизно 12 циклів у рік) на обох графіках зовсім не простежується. Чому? Для цього існують, на наш погляд, як історичні (соціальні) причини, так і природні цикли функціонування організму людини.

Цикл тривалістю приблизно 28 днів можна вважати природнім. Відомі основні періоди місячних циклів: синодичний місяць (проміжок часу між двома фазами Луни) – приблизно 29,5 діб та аномалістичний місяць (проміжок часу між положеннями Луни на орбіті навколо Землі у перигеї або апогеї) – приблизно 27,5 діб. У середньому ці два періоди як раз і дають умовний місяць у 28 днів. Даний період співпадає з середньою тривалістю менструального циклу у жінок, а чоловіки, мабуть, підтягнулись: тому пік у 28 днів (13 циклів на рік) виглядає абсолютно логічно.

Інший виявлений пік у 40 днів важко пояснити природними явищами. Втім, закономірності трафіку, які, зрозуміло, корелюються з біологічними циклами людини та соціальними циклами чомусь вказують, що така «прихована» періодичність існує і якимось впливає на людське суспільство. Причини наявності даного циклу залишимо фахівцям з небесної механіки та фізіології людини. Лише зауважимо, що саме цикл у 40 днів був у календарі древніх слов'ян [26], а також нагадаємо, що день Піднесення християни відмічають на 40-й день після Пасхи. І ще одна (сумна) традиція: другий раз поминати близьких, що пішли, на 40-й день.

Виконаний аналіз дає ще один цікавий факт: у країнах, близьких до екватору, піврічні цикли не спостерігаються на відміну від країн північної півкулі. Дійсно, найкрупніші міста Бразилії (Сан-Пауло та Ріо-де-Жанейро) знаходяться приблизно на 23⁰ південної широти, тоді як Берлін – вже на 52,31⁰ північної широти. Особливості клімату викликають і особливості навантаження на мережі передачі даних: міське населення Європи влітку відпочиває, взимку працює в посиленому режимі або просиджує осінні та весняні дощі за комп'ютером, а сільське населення влітку працює з особливим навантаженням. Що тім, що другим

влітку не до Інтернету. Поряд з цим, у екваторіальних країнах такого різкого сезонного завантаження мережі не спостерігається: банани та кофе плодоносять круглий рік і відпочити людям немає часу, оскільки ніякого літа там не існує, як і зими.

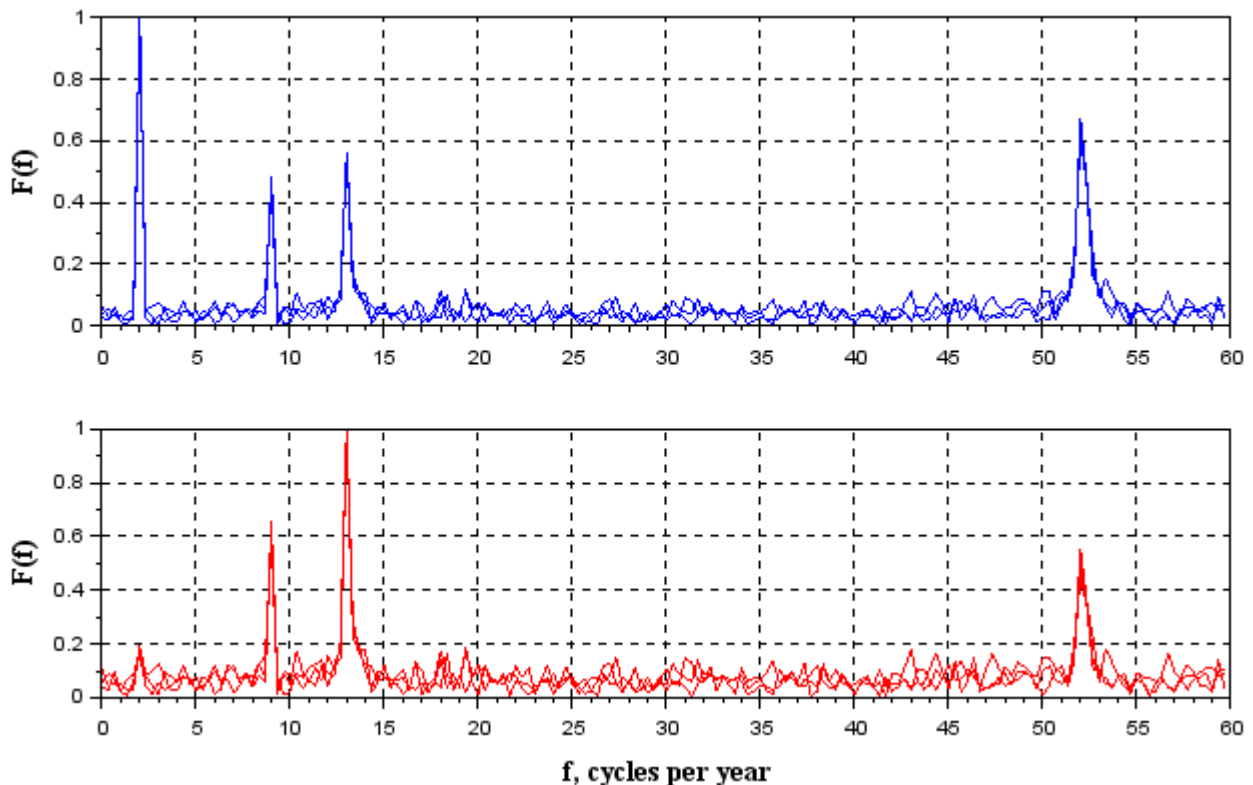


Рис. 9. Порівняння спектрів трафіку у Берліні (зверху) та Бразилії (знизу) на протязі досліджень за останні три роки

Оскільки дослідження у даній роботі продовжують ті, що було розпочато ще у 2019 році [21, 22], то у даному випадку реалізовано найважливіший етап досліджень: перевірку адекватності моделей та ефективності алгоритмів на додаткових наборах даних. Оскільки в основному раніше отримані результати підтверджуються, то можна вважати запропоновану методика досить ефективною, а моделі досить коректними. А саме важливе: за чотири роки у циклах життя людини принципово нічого не змінилось. Про це образно сказав Леонардо да Вінчі в одній зі своїх байок: *«Істина була єдиною дочкою часу»*.

Цікаво було б дослідити цикли у житті людини років так 20000 – 40000 тому методом екстраполяції назад. Чи підтверджувались би тоді визначені періодичності? Нажаль, Інтернету тоді ще не було...

Висновки

У роботі дано основи методики структурної ідентифікації моделей нестационарних квазіперіодичних процесів, яка включає етапи (кроки) неформального фахового аналізу, формалізації задачі та моделі, імітаційного моделювання, застосування розроблених алгоритмів і програм для вирішення прикладної задачі та аналізу адекватності моделі і ефективності алгоритмів на незалежних масивах даних.

Особливістю методики є попереднє диференціювання вхідного процесу з метою пригнічення повільного тренду та визначення слабо виражених «прихованих» компонентів у високочастотному діапазоні. Дана операція сприяє розпізнаванню прихованих мод квазіперіодів та налаштуванню фільтрів нижніх частот для селекції регулярних складових процесу на фоні шумів незалежно від їх особливостей (адитивні, мультиплікативні, фазові).

Застосування даної методики до задачі ідентифікації моделей трафіку на великих вузлах обміну дає очевидні результати наявності періодичностей:

- добовий цикл;
- річний цикл;
- менш очевидних періодичностей:
- піврічний цикл;
- тижневий цикл;
- та зовсім неочевидних періодичностей:
- цикл у 28 днів;
- цикл у 40 днів.

В цілому, запропонована методика дозволяє ефективно вирішувати задачі методики структурної

ідентифікації моделей нестационарних квазіперіодичних процесів при певних обмеженнях:

- основний тренд процесу має бути повільною функцією, похідні якої зменшуються до незначних величин в результаті операції диференціювання;
- характер тренду (приблизна функція, що його моделює) не повинен змінюватись на інтервалі спостережень;
- спектри квазіперіодичних компонентів повинні мати виражені моди та бути сконцентровані у відносно вузьких полосах;
- квазіперіодична складова з мінімальною частотою, що може бути виявлена, повинна мати як найменше 3 періоди за час спостережень;
- квазіперіодична складова з максимальною частотою, що може бути надійно виявлена, повинна бути представлена не менш як 8-ю відліками.

Для виконання моделювання та розрахунків у даній роботі було використано платформу Scilab, яка відрізняється від платформи MatLab принципово тим, що доступна безкоштовно.

Література

1. Харкевич А.А. Спектры и анализ / А.А. Харкевич. – М.: ЛИБКОМ, 2009. – 240 с.
2. A. Golovizin, D. Tregubov, M. Yaushev, D. Mishin, D. Provorchenko, N. Kolachevsky. Absolute frequencies of cooling and pumping transitions for Thulium optical lattice clock and cold gas studies. *Journal of Quantitative Spectroscopy and Radiative Transfer*, Vol. 306, 2023, 108629, ISSN 0022-4073. DOI: <https://doi.org/10.1016/j.jqsrt.2023.108629>.
3. Wei Xie, Guanwen Huang, Le Wang, Pingli Li, Bobin Cui, Haohao Wang, Yu Cao. Long-term performance detection and evaluation of GLONASS onboard satellite clocks. *Measurement*, Volume 175, 2021, 109091, ISSN 0263-2241. DOI: <https://doi.org/10.1016/j.measurement.2021.109091>.
4. J. P. Sánchez Arias, Orlagh L. Creevey. Light curves and frequency analysis of hybrid SCT- dor stars. *Astronomical Journal of Azerbaijan*, 2022, Vol. 17, No. 2, P. 100 – 108.
5. Miruo Lin, Kelai Xi, Yingchang Cao, Rukai Zhu, Xiaobing Niu, Honggang Xin, Weijiao Ma. Cyclicity related to solar activity in lacustrine organic-rich shales and their significance to shale-oil reservoir formation. *Geoscience Frontiers*, Vol. 14, Issue 5, 2023, 101586, ISSN 1674-9871. DOI: <https://doi.org/10.1016/j.gsf.2023.101586>.
6. Tue M. Vu, Ashok K. Mishra. Nonstationary frequency analysis of the recent extreme precipitation events in the United States. *Journal of Hydrology*, Vol. 575, 2019, P. 999-1010, ISSN 0022-1694. DOI: <https://doi.org/10.1016/j.jhydrol.2019.05.090>.
7. Liang-Yu Shyu, Yuh-Shii Lin, Chun-Peng Liu, Wei-Chih Hu. The detection of impedance cardiogram characteristic points using wavelet transform. *Computers in Biology and Medicine*, Volume 34, Issue 2, 2004, P. 165-175, ISSN 0010-4825. DOI: [https://doi.org/10.1016/S0010-4825\(03\)00040-4](https://doi.org/10.1016/S0010-4825(03)00040-4).
8. Jean-Philippe Lachaux, Mario Chavez, Antoine Lutz/ A simple measure of correlation across time, frequency and space between continuous brain signals. *Journal of Neuroscience Methods*, Volume 123, Issue 2, 2003, P. 175-188, ISSN 0165-0270. DOI: [https://doi.org/10.1016/S0165-0270\(02\)00358-8](https://doi.org/10.1016/S0165-0270(02)00358-8).
9. Варакин Л.Е. Теория сложных сигналов. – М.: Сов. радио, 1970. – 376 с.
10. Поллард Дж. Справочник по вычислительным методам статистики / Дж. Поллард. – М.: Финансы и статистика, 1982. – 344 с.
11. Ивахненко А.Г., Степашко В.С. Помехоустойчивость моделирования / А.Г. Ивахненко, В.С. Степашко. – Киев: Наукова думка, 1985. – 216 с.
12. Левин С. Ф. Теория стохастической компактности / С. Ф. Левин // Статистический анализ и синтез систем технического обеспечения эксплуатации. – Одесса: ОВВКИУ ПВО, 1984. – С. 40-178.
13. Залманзон Л.А. Преобразования Фурье, Уолша, Хаара и их применение в управлении, связи и других областях / Л.А. Залманзон. – М.: Наука, 1989. – 496 с.
14. Оппенгейм А.В., Шафер Р.В. Цифровая обработка сигналов / А.В.Оппенгейм, Р.В.Шафер. – М.: Связь, 1979. – 426 с.
15. Рибальченко М.О., Єгоров О.П., Зворикін В.Б. Цифрова обробка сигналів. Навчальний посібник. – Дніпро: НМетАУ, 2018. – 79 с.
16. Воскобойников Ю.Е., Гочаков А.В., Колкер А.Б. Фильтрация сигналов и изображений: Фурье и вейвлет алгоритмы (с примерами в Mathcad) / Ю.Е. Воскобойников, А.В. Гочаков, А.Б. Колкер. – Новосибирск : НГАСУ, 2010. – 188 с.
17. Марченко О.О., Россада Т.В. Актуальні проблеми Data Mining: Навчальний посібник для студентів факультету комп'ютерних наук та кібернетики / О.О. Марченко, Т.В. Россада. – Київ: КНУ ім. Тараса Шевченка, 2017. – 150 с.
18. Анісімов А. В., Дорошенко А. Ю., Погорілий С. Д., Дорогий Я. Ю. Програмування числових методів мовою Python : підруч. / А. В. Анісімов, А. Ю. Дорошенко, С. Д. Погорілий, Дорогий Я. Ю. – К.: Видавничо-поліграфічний центр "Київський університет", 2014. – 640 с.
19. Дьяконов, В.П. Scilab: полный самоучитель / В.П. Дьяконов. – М.: ДМК Пресс, 2012. – 768 с.
20. Чубукова І.О. Data Mining / І.О. Чубукова. Київ: КНЕУ ім. В. Гетьмана, 2016. – 326 с.

[електронний ресурс] – Режим доступу: http://kist.ntu.edu.ua/textPhD/Chubukova-Data_Mining.pdf.

21. Odegov N.A. Analysis of non-stationary quasi-periodic internet traffic by the method of instant spectra / Odegov N.A., Bahachuk D.H., Hordiienko O.O. // Наукові праці ОНАЗ ім. О.С. Попова, 2019, № 1. – С. 121 – 134.

22. Одегов М.А. Метод селекції гармонійних складових нестационарної моделі трафіку на великих вузлах обміну / Одегов М.А., Горієднко О.О., Юр'єва О.В. // Матеріали XIV міжнародної наук.-техн. конференції "Вимірювальна та обчислювальна техніка в технологічних процесах" (ВОТТП-19), 14-19 червня 2019 р., Одеса (Затока). – С. 47 – 49.

23. Статистика трафіку у бразильській точці обміну [електронний ресурс] – Режим доступу: <https://ix.br/agregado/>.

24. Статистика трафіку у берлінській точці обміну [електронний ресурс] – Режим доступу: <https://www.bcix.de/bcix/traffic/>.

25. Гоноровский И.С. Радиотехнические цепи и сигналы / И.С. Гоноровский. – М.: Радио и связь, 1986. – 512 с.

26. Славянский календарь: месяц 40 дней, неделя 9 дней [Электронный ресурс]. – Режим доступа: <http://derzhavarus.ru/kalendar-let-mesyac-nedelya.html>.

References

1. Harkevich A.A. Spektry i analiz / A.A. Harkevich. – М.: LIBKOM, 2009. – 240 с.
2. A. Golovizin, D. Tregubov, M. Yaushev, D. Mishin, D. Provorchenko, N. Kolachevsky. Absolute frequencies of cooling and pumping transitions for Thulium optical lattice clock and cold gas studies. Journal of Quantitative Spectroscopy and Radiative Transfer, Vol. 306, 2023, 108629, ISSN 0022-4073. DOI: <https://doi.org/10.1016/j.jqsrt.2023.108629>.
3. Wei Xie, Guanwen Huang, Le Wang, Pingli Li, Bobin Cui, Haoao Wang, Yu Cao. Long-term performance detection and evaluation of GLONASS onboard satellite clocks. Measurement, Volume 175, 2021, 109091, ISSN 0263-2241. DOI: <https://doi.org/10.1016/j.measurement.2021.109091>.
4. J. P. Sanchez Arias, Orlagh L. Creevey. Light curves and frequency analysis of hybrid SCT- dor stars. Astronomical Journal of Azerbaijan, 2022, Vol. 17, No. 2, R. 100 – 108.
5. Miruo Lin, Kelai Xi, Yingchang Cao, Rukai Zhu, Xiaobing Niu, Honggang Xin, Weijiao Ma. Cyclicity related to solar activity in lacustrine organic-rich shales and their significance to shale-oil reservoir formation. Geoscience Frontiers, Vol. 14, Issue 5, 2023, 101586, ISSN 1674-9871. DOI: <https://doi.org/10.1016/j.gsf.2023.101586>.
6. Tue M. Vu, Ashok K. Mishra. Nonstationary frequency analysis of the recent extreme precipitation events in the United States. Journal of Hydrology, Vol. 575, 2019, P. 999-1010, ISSN 0022-1694. DOI: <https://doi.org/10.1016/j.jhydrol.2019.05.090>.
7. Liang-Yu Shyu, Yuh-Shii Lin, Chun-Peng Liu, Wei-Chih Hu. The detection of impedance cardiogram characteristic points using wavelet transform. Computers in Biology and Medicine, Volume 34, Issue 2, 2004, P. 165-175, ISSN 0010-4825. DOI: [https://doi.org/10.1016/S0010-4825\(03\)00040-4](https://doi.org/10.1016/S0010-4825(03)00040-4).
8. Jean-Philippe Lachaux, Mario Chavez, Antoine Lutz/ A simple measure of correlation across time, frequency and space between continuous brain signals. Journal of Neuroscience Methods, Volume 123, Issue 2, 2003, P. 175-188, ISSN 0165-0270. DOI: [https://doi.org/10.1016/S0165-0270\(02\)00358-8](https://doi.org/10.1016/S0165-0270(02)00358-8).
9. Varakin L.E. Teoriya slozhnykh signalov. – М.: Sov. radio, 1970. – 376 с.
10. Pollard Dzh. Spravochnik po vychislitel'nykh metodam statistiki / Dzh. Pollard. – М.: Finansy i statistika, 1982. – 344 с.
11. Ivahnenko A.G., Stepashko V.S. Pomehoustojchivost modelirovaniya / A.G. Ivahnenko, V.S. Stepashko. – Kiev: Naukova dumka, 1985. – 216 с.
12. Levin S. F. Teoriya stohasticheskoy kompaktnosti / S. F. Levin // Statisticheskij analiz i sintez sistem tehničeskogo obespecheniya ekspluatatsii. – Odessa: OVVKIU PVO, 1984. – С. 40-178.
13. Zalmanzon L.A. Preobrazovaniya Fure, Uolsha, Haara i ih primenenie v upravlenii, svyazi i drugih oblastyah / L.A. Zalmanzon. – М.: Nauka, 1989. – 496 с.
14. Oppengejm A.V., Shafer R.V. Cifrovaya obrabotka signalov / A.V.Oppengejm, R.V.Shafer. – М.: Svyaz, 1979. – 426 с.
15. Ribalchenko M.O., Yegorov O.P., Zvorikin V.B. Cifrova obrobka signaliv. Navchalnij posibnik. – Dnipro: NMetAU, 2018. – 79 с.
16. Voskobojnikov Yu.E., Gochakov A.V., Kolker A.B. filtraciya signalov i izobrazhenij: Fure i vejjvlet algoritmy (s primerami v Mathcad) / Yu.E. Voskobojnikov, A.V. Gochakov, A.B. Kolker. – Novosibirsk : NGASU, 2010. – 188 с.
17. Marchenko O.O., Rossada T.V. Aktualni problemi Data Mining: Navchalnij posibnik dlya studentiv fakultetu komp'yuternih nauk ta kibernetiki / O.O. Marchenko, T.V. Rossada. – Kiyiv: KNU im. Tarasa Shevchenka, 2017. – 150 с.
18. Anisimov A. V., Doroshenko A. Yu., Pogorilij S. D., Dorogij Ya. Yu. Programuvannya chislovih metodiv movoyu Python : pidruch. / A. V. Anisimov, A. Yu. Doroshenko, S. D. Pogorilij, Dorogij Ya. Yu. – K.: Vidavnicno-poligrafichnij centr "Kiyivskij universitet", 2014. – 640 с.
19. Dyakonov, V.P. Scilab: polnyj samouchitel / V.P. Dyakonov. – М.: DMK Press, 2012. – 768 с.
20. Chubukova I.O. Data Mining / I.O. Chubukova. Kiyiv: KNEU im. V. Getmana, 2016. – 326 с. [електронний ресурс] – Rezhim dostupu: http://kist.ntu.edu.ua/textPhD/Chubukova-Data_Mining.pdf.
21. Odegov N.A. Analysis of non-stationary quasi-periodic internet traffic by the method of instant spectra / Odegov N.A., Bahachuk D.H., Hordiienko O.O.// Naukovyi pracj ONAZ im. O.S. Popova, 2019, № 1. – С. 121 – 134.
22. Odegov M.A. Metod selekciji harmonijnih skladovih nestacionarnej modeli trafiku na velikih vuzlah obminu / Odegov M.A., Goriyednko O.O., Yur'yeva O.V. // Materiali XIV mizhnarodnoyi nauk.-tehn. konferenciji "Vimiryuvalna ta obchislyuvalna tehnika v tehnologichnih procesah" (VOTTP-19), 14-19 chervnya 2019 r., Odessa (Zatoka). – С. 47 – 49.
23. Statistika trafiku u brazil'skij tochci obminu [elektronnij resurs] – Rezhim dostupu: <https://ix.br/agregado/>.
24. Statistika trafiku u berlinskij tochci obminu [elektronnij resurs] – Rezhim dostupu: <https://www.bcix.de/bcix/traffic/>.
25. Gonorovskij I.S. Radiotekhnicheskie cepi i signaly / I.S. Gonorovskij. – М.: Radio i svyaz, 1986. – 512 с.
26. Slavjanskij kalendar: mesyac 40 dnei, nedelya 9 dnei [Elektronnyj resurs]. – Rezhim dostupa: <http://derzhavarus.ru/kalendar-let-mesyac-nedelya.html>.