

**К.т.н., доцент Маслянюк П.П., аспірант Землянський Ю.Р.**

**Національний технічний університет України  
«Київський політехнічний інститут»**

## **ФРАКТАЛЬНА МОДЕЛЬ ТА РЕАЛІЗАЦІЯ АЛГОРИТМУ АНАЛІЗУ ЧИСЛОВИХ РЯДІВ**

### **Вступ**

Прогнозування застосовується у багатьох галузях починаючи з менеджменту і закінчуючи прогнозом погоди. Однією із найактуальніших сфер, де використовуються методи прогнозування, є фондові ринки. Оскільки, знаючи тенденцію поведінки ринку, можна зменшити ризики втрати великої кількості коштів та забезпечити більший прибуток. Так, наприклад, щоденний обсяг торгів на біржах США перевищує 50 мільярдів доларів, а щоденний обсяг купівлі і продажу середньої американської компанії становить близько 3-х мільйонів доларів. Оскільки складну систему складно або неможливо змодельювати, то аналіз її роботи і прогнозування майбутніх результатів системи зводиться до аналізу часового ряду, одному з результатів роботи системи.

Тому проблема підбору моделі, яка б найбільш адекватно описувала часовий ряд, є досить актуальною.

Одними з найкращих та розповсюджених на даний час є моделі (різні модифікації GARCH (generalized autoregressive conditional heteroskedasticity)), які базуються на принципі застосованому в ARMA (autoregressive moving average model). Однак ці моделі не враховують фрактальної структури даних.

Тому ця робота присвячена аналізу доцільності застосування теорії фракталів та можливості використання її для створення засобів збільшення адекватності моделей прогнозування.

### **Постановка задачі**

Головними задачами роботи є:

- Дослідити і проаналізувати можливість та доцільність використання теорії фракталів;
- Розробити компонентну структуру системи прогнозування;
- Реалізувати алгоритм аналізу часових рядів на фрактальність.

## Методи дослідження числових рядів на фрактальність

Застосування системного підходу для аналізу моделей прогнозування дозволяє абстрагуватися від деталей реалізації кожної моделі та виділити основні загальносистемні критерії їх функціонування. Хронологічний розвиток моделей прогнозування можна охарактеризувати виявленням нових параметрів, які характеризують поведінку часового ряду, та створенням і розвитком моделей, що використовують ці параметри. Тим самим вдосконалюючи існуючу систему прогнозування (СП) і наближаючи її до «ідеальної» СП.

Під СП ми розуміємо уявну систему, яка складається з усіх моделей для прогнозування та їх реалізацій і яка розширює свої можливості з появою нових моделей. Нові моделі, в свою чергу, можуть бути, як моделями, похідними від існуючих, так і моделями, створеними за рахунок використання нових параметрів, які характеризують поведінку часових рядів. Надалі термін «параметри часового ряду» ми будемо розуміти саме в такому контексті.

Одним з найпростіших є метод прогнозування, що ґрунтується на аналізі відхилення даних від середнього значення - математичного сподівання. Наступною ознакою, принципом, який використовувався, було твердження, що майбутні значення ряду найкраще характеризуються не всією вибіркою даних, а лише певним об'ємом (за часовою шкалою). При такому підході першими з'являються моделі аналогічні «ковзаючому середньому».

Для збільшення ефективності моделі адаптують, будуючи їх таким чином, що чим старіші дані, тим вони менше повинні впливати на майбутній результат. Для цього використовують експоненціальне згладжування та аналіз тренду, з'являється модель Хольта і Брауна.

Наступним етапом в розвитку системи прогнозування є підхід, якій враховує періодичність коливань даних. Разом з цим з'являються регресивні моделі, які можна охарактеризувати тим, що вони доповнюють СП моделями, які використовують поняття незалежних та залежних змінних, чинників.

Далі з'являються методики і моделі, які самі адаптують свої параметри (параметри моделі) для поліпшення адекватності - моделі Бокса-Дженкінса [1] (ARMA), GARCH та ін.

Разом з цим розвивається напрямок, який використовує штучні нейронні мережі (ШНМ) для прогнозування. Така техніка підходу не виділяється використанням нових параметрів ряду, а лише використовує вже перелічені методики для аналізу та фільтрації вхідних даних, в процесі навчання та прогнозування за допомогою ШНМ.

Тому дослідження використання теорії фракталів [2] для створення і реалізації моделей прогнозування, які б могли використовувати нові параметри часового ряду для поліпшення результатів прогнозування, є перспективними і необхідними.

Існує декілька методів для оцінки часового ряду з точки зору наявності ознак фрактальності. Один із універсальних підходів з виявлення самоподібності ґрунтується на методі DFA (Detrended Fluctuation Analysis) [3] – універсальному методі обробки часових рядів. Цей підхід являє собою варіант дисперсійного аналізу, який дозволяє досліджувати ефекти довгих кореляцій в нестационарних рядах. Для більш скрупульозного дослідження скейлінгу ряду застосовується метод максимумів модулів вейвлет-перетворення (ММВП [4]).

### **Компонентна модель системи прогнозування**

Для реалізації системи використано підхід, в основі якого система ділиться на компоненти на основі певних загальносистемних ознак [5]. Для графічного представлення моделі системи використано UML. Компонентна модель системи прогнозування (рис. 1) складається з компонентів та інтерфейсів, що реалізують взаємодію між ними. Вона забезпечує масштабованість і розширюваність проекту від ітерації до ітерації та незалежність процесів розробки окремих компонентів. В моделі представлені компоненти, які виконують функції прогнозування, управління, аналітичної обробки прогнозованих даних та методів. В свою чергу компоненти для прогнозування поділяються на три типи, які використовують стандартні моделі, технологію ШНМ, та моделі які використовують теорію фракталів.

### **Реалізація алгоритму аналізу числових рядів**

На рис. 2 представлено модель визначення фрактальних характеристик часового ряду. В ній зображено діаграму діяльності і взаємодії двох компонент «Фрактальний аналіз» та «Робота з даними». Для реалізації аналізу використовуються методи такі, як: detrended fluctuation analysis (DFA), метод визначення фактору Фано, аналіз показника Херста, та метод максимумів вейвлет-перетворення (ММВП). В рамках алгоритму DFA аналізується середньоквадратична похибка лінійної апроксимації в залежності від розміру апроксимованої ділянки. Спочатку виконується зведення даних рядів релеванностей до нульового середнього і побудови випадкового блукання  $y(k)$ . Після цього ряд значень  $y(k)$  розбивається на відрізки, що не перекриваються, в межах кожного з яких визначається

рівняння прямої, що апроксимує послідовність  $y(k)$ . Далі підраховується середньоквадратична похибка лінійної апроксимації  $F(n)$  та відповідні розрахунки проводяться в широкому діапазоні значень  $n$ . Вважається, що залежність  $F(n)$  часто має степеневий характер  $F(n) \sim n^\alpha$ , а наявність лінійної ділянки в подвійному логарифмічному масштабі  $\lg F(\lg n)$  дозволяє говорити про існування скейлінгу.

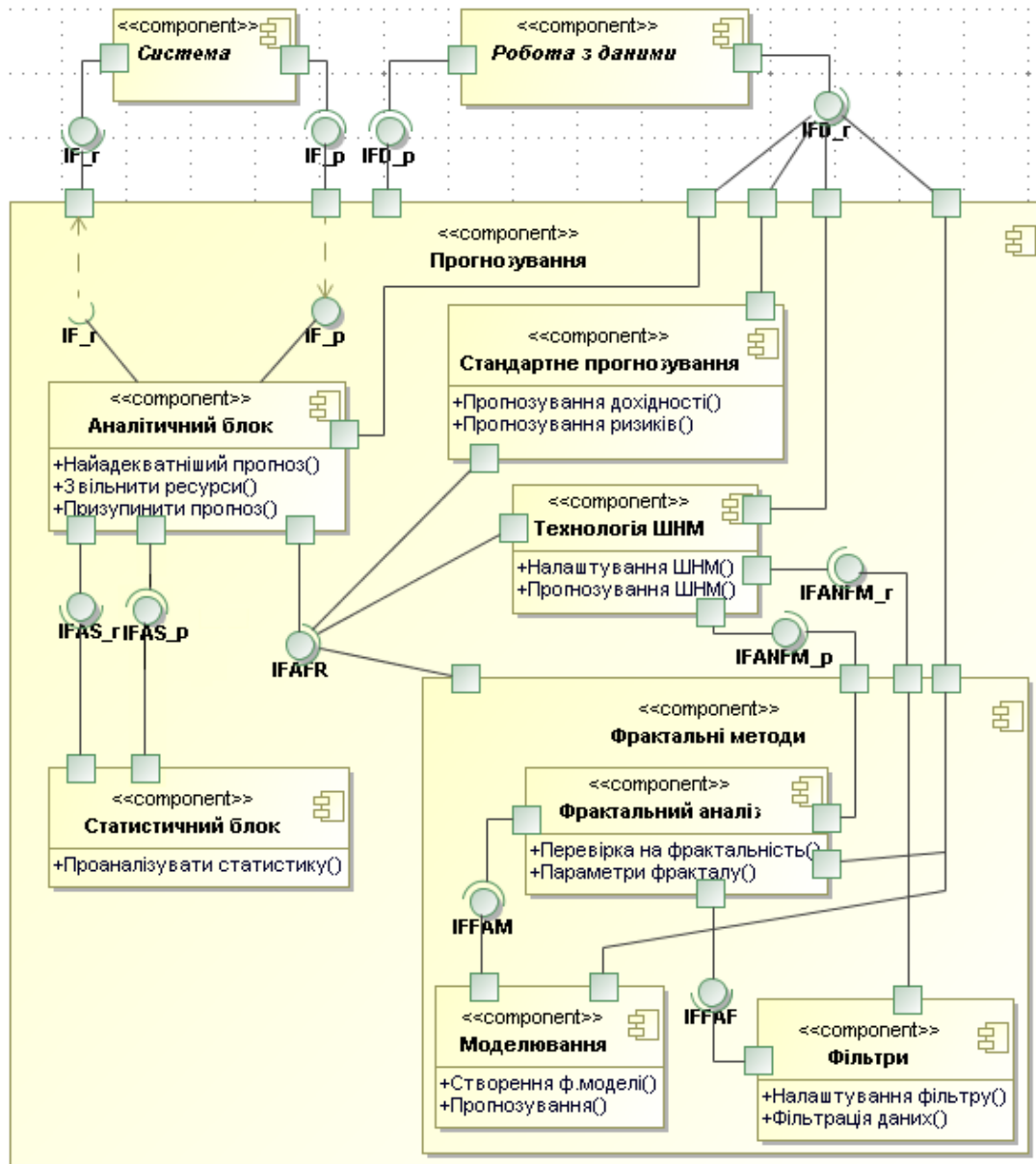


Рис. 1. Модель прогнозування. Діаграма компонентів в нотації UML

На практиці величина  $\alpha$  (називається скейлінговою експонентою DFA-методу) може відрізнитись для різних  $n$ , що свідчить про зміну

властивостей скейлінгу при збільшенні масштабу. В такій ситуації є доцільним проводити аналіз локальних показників  $\alpha$ .

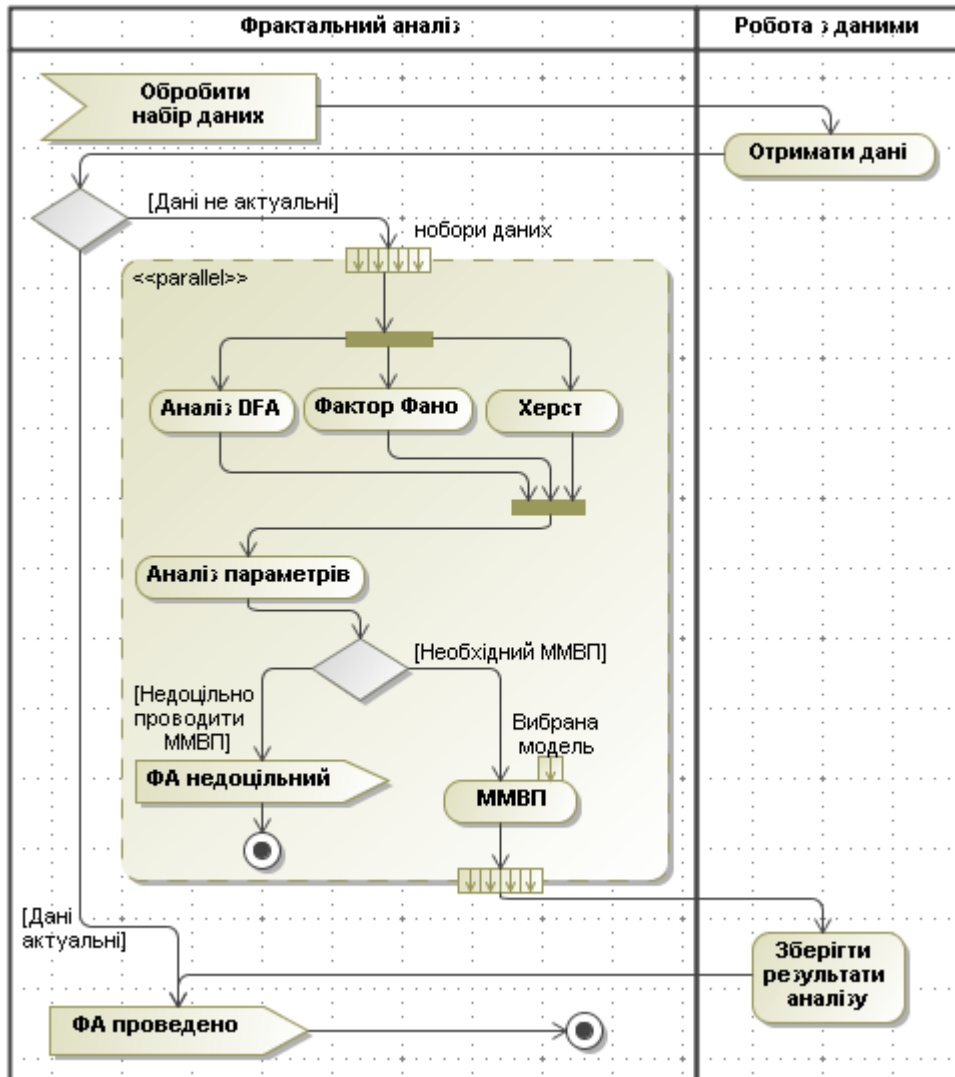


Рис. 2. Модель визначення фрактальних характеристик часових рядів.  
 Діаграма діяльності в нотатції UML

Фактор Фано використовується для підтвердження самоподібності числових рядів, це є не що інше як індекс розкиду дисперсії (ІРД). Ця величина ( $F(k)$ ) визначається як відношення дисперсії досліджуваного числового ряду на заданому вікні спостережень  $k$  до відповідного математичного очікування. Для самоподібних числових рядів виконується функція  $F(k) \cong Ck^\alpha$ , де  $C$  і  $\alpha$  – константи. При цьому у випадку фрактальної структури числових рядів виконується співвідношення  $\alpha = 2H - 1$ , де  $H$  – показник Херста, безпосередньо зв'язаний з фрактальною розмірністю. Показник Херста ( $H$ ) зв'язують з коефіцієнтом нормованого розмаху

( $R/S$ ), де  $R$  – обчислений певним чином «розмах» відповідного часового ряду, а  $S$  – стандартне відхилення.

На основі проведених досліджень, можна визначити необхідність та доцільність аналізу ММВП. В рамках якого спочатку часовий ряд розкладається, за допомогою вейвлет-перетворення. І далі проводиться аналіз ліній екстремумів вейвлет-перетворення.

## Висновки

В результаті проведеної роботи встановлено, що для часових рядів, які характеризують показники фінансових ринків, характерне проявлення мультифрактальних ознак. В ході аналізу існуючих методів прогнозування виявлено, що вони не враховують мультифрактальні характеристики, які є в числових рядах. Причому, рівень прояву мультифрактальних характеристик в різних числових рядах, різний.

Таким чином можна стверджувати, що використання теорії фракталів для прогнозування показників фінансових ринків є актуальним і перспективним напрямком дослідження.

Одним із напрямків подальших досліджень є створення моделі прогнозування, яка б враховувала фрактальну структуру часового ряду. Також перспективним напрямком роботи є подальша реалізація всіх компонентів запропонованої моделі системи.

## Література

1. *George Box, Gwilym M. Jenkins, and Gregory C. Reinsel. Time Series Analysis: Forecasting and Control, third edition. Prentice-Hall, 1994. - 746с.*
2. *E. Bacry, A. Kozhemyak, J.-F. Muzy Log-Normal continuous cascades: aggregation properties and estimation. Application to financial time-series, 2008. - 27с. (<http://arxiv.org/abs/0804.0185>).*
3. *Снарский А. А., Ландэ Д. В., Брайчевский С. М., Дармохвал А. Т. “Распределение документов по степени релевантности на основе мультифрактальных свойств”. - 10 с. (<http://company.yandex.ru/academic/grant/report2007.xml>).*
4. *A.N. Pavlov, V.S. Anishchenko Multifractal analysis of complex signals International Research Institute of Nonlinear Dynamics, N.G. Chernyshevskii Saratov State University 2007. - С.859-876.*
5. *Масляно П.П. Компонентні процеси системного проектування інформаційно-комунікаційних систем//Наукові вісті НТУУ “КПІ”. 2008, № 2. - С. 112–121.*