

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**  
**КРИВОРІЗЬКИЙ ДЕРЖАВНИЙ ПЕДАГОГІЧНИЙ УНІВЕРСИТЕТ**  
Фізико-математичний факультет  
Кафедра інформатики та прикладної математики

«Допущено до захисту»  
В.о. завідувача кафедри  
\_\_\_\_\_ Моїсеєнко Н. В.  
« \_\_\_ » \_\_\_\_\_ 2023 р.

Реєстраційний № \_\_\_\_\_  
« \_\_\_ » \_\_\_\_\_ 20\_\_ р.

**МАСШТАБНО-ІНВАРІАНТНИЙ ЕНТРОПІЙНИЙ АНАЛІЗ**  
**СКЛАДНИХ СИСТЕМ**

Кваліфікаційна робота студента групи І-19  
ступінь вищої освіти бакалавр  
спеціальності

014.09 Середня освіта (Інформатика)

**Рябікова Максима Олеговича**

Керівник

**професор, д. ф.-м. н. Соловйов В. М.**

Оцінка:

Національна шкала \_\_\_\_\_

Шкала ECTS \_\_\_\_\_ Кількість балів \_\_\_\_\_

Голова ЕК \_\_\_\_\_

Члени ЕК \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

Кривий Ріг

2023

## **ЗАПЕВНЕННЯ**

Я, Рябіков Максим Олегович, розумію і підтримую політику Криворізького державного педагогічного університету з академічної доброчесності. Запевняю, що ця кваліфікаційна робота виконана самостійно, не містить академічного плагіату, фабрикації, фальсифікації. Я не надавав і не одержував недозволену допомогу під час підготовки цієї роботи. Використання ідей, результатів і текстів інших авторів мають покликання на відповідне джерело. Із чинним Положенням про запобігання та виявлення академічного плагіату в роботах здобувачів вищої освіти Криворізького державного педагогічного університету ознайомлений. Чітко усвідомлюю, що в разі виявлення у кваліфікаційній роботі порушення академічної доброчесності робота не допускається до захисту або оцінюється незадовільно.

## ЗМІСТ

ВСТУП .....	4
РОЗДІЛ 1. РИНОК БІТКОЇНА. АНАЛІЗ БІБЛІОГРАФІЇ	
ПРИСВЯЧЕНУ РИНКУ БІТКОЇНА.....	6
1.1 Що таке біткоїн і чи можна ним користуватись? .....	6
1.2 Справжня вартість біткоїна на ринку .....	8
1.4 Висновки до розділу 1 .....	12
РОЗДІЛ 2. МОЖЛИВІСТЬ ВИКОРИСТАННЯ ІНСТРУМЕНТАРІЮ	
ЕНТРОПІЙНИХ ПОКАЗНИКІВ СКЛАДНОСТІ .....	13
2.1 Ентропія. Апроксимаційна ентропія та ентропія шаблонів .....	13
2.2 Пермутаційна ентропія: природна міра складності часових рядів	17
2.3 Дисперсійна ентропія. Міра для аналізу часових рядів.....	19
2.4 Нечітка ентропія. Вимірювання регулярності часових рядів за допомогою нелінійної вибірки на основі подібності ентропії. ....	22
2.5 Висновки до розділу 2.....	24
РОЗДІЛ 3. ПОРІВНЯЛЬНИЙ АНАЛІЗ ПРЕДСТАВЛЕНИХ	
ПОКАЗНИКІВ, ОПИРАЮЧИСЬ НА ЧАСОВИЙ РЯД БІТКОЇНА.....	24
3.1 Апроксимаційна ентропія .....	26
3.2 Дисперсійна ентропія .....	28
3.3 Нечітка ентропія.....	30
3.4 Пермутаційна ентропія.....	32
3.5 Ентропія шаблонів .....	34
ВИСНОВКИ.....	37
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	38

## ВСТУП

Значна кількість досліджень була присвячена поясненню зміни в ціні біткоїна і прогнозування його майбутньої динаміки. Хоча ці дослідження мали успіх при прогнозуванні, опираючись на відносно стабільну цінову динаміку, жодне з них не мало успіху при передчасній ідентифікації значної волатильності цін на біткоїн. Останні дослідження досягли більшого успіху, використовуючи авторегресивні моделі, вейвлет аналіз та нейронні мережі, щоб пояснити рух цін на біткоїн, принаймні в короткостроковій перспективі, на основі минулих тенденцій та настроїв інвесторів. Цикли зворотного зв'язку та механізми, що фіксують ці моделі, добре підходять для пояснення волатильності цін на біткоїн, але не передбачають довгострокової поведінки біткоїна, оскільки вони не включають складність даного активу в повній мірі.

**Актуальність теми.** Біткоїн (BTC) – це криптовалюта, віртуальна валюта, створена для того, щоб діяти як гроші та форма платежу поза контролем будь-якої особи, групи чи організації, таким чином усуваючи потребу в участі третіх сторін у фінансових операціях. Він нагороджується майнерами блокчейнів за виконану роботу з перевірки транзакцій, і його можна придбати на кількох біржах.

Біткоїн був представлений громадськості в 2009 році анонімним розробником або групою розробників на ім'я Сатоші Накамото.

З тих пір вона стала найвідомішою криптовалютою в світі. Його популярність надихнула на розвиток багатьох інших криптовалют. Ці конкуренти або намагаються замінити її як платіжну систему, або використовуються як службові або захисні токени в інших блокчейнах і нових фінансових технологіях.

У серпні 2008 року було зареєстровано доменне ім'я Bitcoin.org. Принаймні сьогодні цей домен захищено WhoisGuard, тобто інформація про особу, яка його зареєструвала, не є публічною інформацією.

У жовтні 2008 року особа або група під псевдонімом Сатоші Накамото оголосила про список розсилки криптографії на metzdowd.com: «Я працював

над новою електронною готівковою системою, яка є повністю одноранговою, без довіреної третьої сторони. Ця нині відома біла книга, опублікована на Bitcoin.org під назвою «Біткоїн: однорангова електронна готівкова система», стане Великою хартією вольностей для того, як сьогодні працює біткоїн.

3 січня 2009 року було видобуто перший блок біткоїн – блок 0. Він також відомий як «блок генезису» і містить текст: «The Times 03/Jan/2009 Chancellor на порозі другої допомоги банкам». можливо, доказ того, що блок було видобуто в цю дату або пізніше, а також як відповідний політичний коментар.

Винагорода в біткоїнах зменшується вдвічі кожні 210 000 блоків. Наприклад, у 2009 році винагорода за блок становила 50 нових біткоїнів. 11 травня 2020 року відбулося третє вдвічення, що призвело до зниження винагороди за кожне відкриття блоку до 6,25 біткоїнів.

**Мета роботи.** Представлення ефективних індикаторів або індикаторів-передвісників кризових явищ криптовалютного ринку на основі класичного інструментарію ентропійних показників складності та їх мультимасштабних аналогів

**Об'єктом дослідження** ринок біткоїна.

**Предметом** слугує методологія та інструментарій класичних ентропійних показників складності та їх мультимасштабних аналогів.

**Завдання дослідження:**

1. Проаналізувати бібліографію присвячену ринку криптовалют та ентропійних показників різних типів.

2. Дослідити можливість використання інструментарію ентропійних показників складності та їх мультимасштабних аналогів у якості індикаторів або індикаторів-передвісників кризових явищ.

3. Провести порівняльний аналіз представлених показників, опираючись на часовий ряд біткоїна.

**Структура та обсяг роботи:** робота складається із вступу, трьох розділів, висновків до кожного розділу, загальних висновків та списку використаних джерел, що нараховує 9 джерел. Обсяг роботи 38 сторінок.

## РОЗДІЛ 1. РИНОК БІТКОЇНА. АНАЛІЗ БІБЛІОГРАФІЇ ПРИСВЯЧЕНУ РИНКУ БІТКОЇНА

### 1.1 Що таке біткоїн і чи можна ним користуватись?

Біткоїн є найбільшим блокчейном і забезпечує базову архітектуру УТХО, яка використовується багатьма іншими криптовалютами.

Критично важливими для аналізу будь-якої системи, заснованої на платежах, є точні показники її прийняття, використання та вартості. Незважаючи на «прозорий» характер блокчейну біткоїн, специфічні особливості його реалізації дають істотно оманливі сигнали про кількість і вартість платежів, а також про справжню кількість користувачів в екосистемі. Розмір і масштаб цих накопичених транзакцій (блокчейн Bitcoin накопичив сотні мільйонів транзакцій за останнє десятиліття) ускладнюють типовим фінансово-економічним дослідникам ізоляцію і кількісну оцінку величини цих завищень.

Надавши основу для виправлення цієї ситуації, створюючи евристику, за допомогою якої всі дослідники можуть визначити справжню вартість транзакцій і витрат в мережі біткоїн (і будь-який інший УТХО), а також точно кількісно оцінити справжню кількість унікальних користувачів. Щоб гарантувати, що величезний розмір даних, з якими стикаються особи, не стане занадто складним, ми також надаємо пов'язаний щоденний набір даних, що включає як очевидні, так і скориговані цифри.

Швидке зростання ціни та ринкової капіталізації криптовалют спричинило відповідне збільшення академічної літератури, яка прагне зрозуміти їх динаміку ціноутворення та забезпечити рамки оцінки. Численні фактори були емпірично перевірені на предмет їх прогностичної здатності в поясненні віддачі від криптовалют. До них відносяться: настрої в Twitter (Kraaijeveld and De Smedt, 2020); регуляторні втручання (Meegan, Corbet,

Larkin and Lucey, 2021); пастуша поведінка (Філіппас, Філіппас, Ціогкідіс і Джиба, 2020); попит на лотерейні активи (Grobys and Junttila, 2021); як оплата в атаках програм-вимагачів (Соколов, 2021; Капорале, Канг, Спаньоло та Спаньоло, 2021); їх здатність створювати анонімні платежі (Сапкота і Гробіс, 2021); їх кореляція з цінами акцій, іноземною валютою та товарами (Borri 2019; Panagiotidis et al. 2019; Дирберг 2016); існування спекулятивного попиту (Лі, Лі та Чжен, 2020; Фолі та ін., 2022); і як розрахункова одиниця на децентралізованих біржах (Aspris et al. 2021; Svec et al. 2020).

Вивчаючи 600 мільйонів адрес біткоїн-гаманців, що включають 430 мільйонів біткоїн-транзакцій, щоб оцінити, наскільки обсяги транзакцій, витрати користувачів та адреси гаманців у протоколі Bitcoin є упередженими та можуть затьмарити наше розуміння прийняття цього конкретного блокчейну, а також приховати справжні витрати на транзакції. Наш новий алгоритм методології ідентифікації транзакцій (TIM) досягає 95% точності при відокремленні «економічної цінності» від «цінності безпеки», щоб забезпечити більш точне представлення реальності, покращуючи прийняття рішень і зменшуючи суттєве спотворення заходів з прийняття.

Результати показують, що унікальний механізм, присутній у блокчейнах Unspent Transaction Output (UXTO), який вимагає від користувачів «витрачати» весь вміст адреси гаманця, щоб здійснити платіж, а потім отримати «зміну» на унікальну нову адресу, значно завищує обсяги транзакцій, заплує витрати на транзакції та завищує кількість унікальних користувачів. Ми надаємо ці щоденні скориговані показники для використання майбутніми вченими в додатку до Інтернету.

Висновки підкреслюють небезпеку екстраполяції рівнів прийняття блокчейнів UXTO, оскільки інфляція номерів транзакцій з автоматично згенерованих рахунків гаманців буде тільки збільшуватися зі зростанням фіатної вартості токена (наприклад, біткоїна), створюючи видимість зростання користувачів, навіть якщо нових користувачів не існує.

## 1.2 Справжня вартість біткоїна на ринку

Аналіз валютного ціноутворення часто використовує моделі, засновані на кількісній теорії грошей (QTM) або паритеті купівельної спроможності (ДПП). Ці моделі мало корисні з біткоїнами, оскільки в даний час біткоїн не є широко поширеною розрахунковою одиницею або засобом обміну (Baig, Hong, & Lee, 2018). Хоча деякі компанії приймають біткоїни, набагато менше коштують їхні продукти та послуги в цій криптовалюті. Замість цього вони використовують фіатні валюти своїх економічних регіонів, такі як долари або євро, щоб оцінити свої товари продажу.

Якщо вони приймають біткоїн, клієнти конвертують ці ціни відповідно до спотового обмінного курсу (тобто ціни біткоїна). Як наслідок, неможливо створити індекс цін на біткоїн, подібний до індексу споживчих цін або неявного індексу цін, який існує для товарів, номінованих у доларах, що споживаються або виробляються в Сполучених Штатах, або для товарів і послуг, номінованих у місцевій валюті будь-якої іншої країни. Моделі QTM і PPP покладаються на такі індекси або їх замітники, але жодна з них не здається надійною для біткоїна.

На конкурентних ринках ціни визначаються силами потоку попиту і пропозиції (Cecchetti & Schoenholtz, 2021; Мартінсен, 2020). Рівноважний обмінний курс знаходиться на перетині низхідно-похилого попиту і висхідно-похилої пропозиції. Як і будь-яка вільно коливається валюта, ціна біткоїна вимірюється вартістю його придбання у фіатних валютах, альткоїнах або дорогоцінних металах, і вона визначається *поток*ом сил попиту та пропозиції за період. Учасники ринку долар-біткоїни потенційно ідентичні учасникам ринку долар-євро, але рівні їх участі і масштаби залучення помітно відрізняються. Переважна більшість покупців біткоїни - це інвестори і спекулянти, яких стимулюють в основному *зміни* прибутковості і *коливання* очікуваних ризиків і прибутковості.



Ринок долар-біткоїни має дві характеристики чисто конкурентного ринку, однорідного активу і відсутності бар'єрів входу або виходу. Кожен, хто бажає купити або продати біткоїни, може зробити це через біржі або безпосередньо за допомогою протоколу біткоїн. Питання про те, чи має ринок біткоїнів останні дві характеристики чисто конкурентного ринку (тобто багато покупців і продавців та ідеальну інформацію), є дискусійним. Ринок відносно неглибокий, а це означає, що покупець або продавець з великим обсягом може значно зрушити ринкову ціну. Більше того, ринкова інформація не є ідеальною, тому в перші дні були можливі розбіжності в цінах між біржами. Останнім часом кращі інформаційні потоки, викликані підвищеним інтересом біткоїна, зменшили або усунули цю здатність.

Майнери на ринку долар-біткоїн є ціноутворювачами, а не ціноутворювачами. В результаті вони стикаються з ідеально еластичними кривими попиту, а це означає, що їх середні та граничні доходи рівні. Якщо середньостатистичний майнер заробляє надприбутки, то нові майнери повинні вийти на ринок і зменшити або усунути ці надлишкові прибутки (Крістуфек, 2020).

Останні емпіричні дослідження підтверджують наш висновок про те, що зміни вартості не мають значного впливу на ціну біткоїна і слідує за змінами цін, а не ведуть їх. Наприклад, Kristoufek (2020), використовуючи модель коінтеграції та виправлення векторних помилок для даних з січня 2014 року по липень 2018 року, робить висновок: «Ціна біткоїна стимулює витрати на майнінг, а не (або лише слабко) навпаки». Фантаціні та Колодін (2020), дані яких тривають з 1 серпня 2015 року по 29 лютого 2020 року, роблять висновок, що «у першому дослідженому зразку [з 1 серпня 2016 року по 4 грудня 2017 року] не було ні доказів причинності Грейнджера, ні коінтеграції, ..., тоді як у другій вибірці були докази односпрямованої причинності Грейнджера та коінтеграції [з 11 грудня 2017 року по 24 лютого, 2020] ..., переходячи від ціни біткоїна до хешрейту (або до [вартості виробничих моделей]), але не навпаки." Kjørland та інші (2018), дані яких тривають з 1 січня 2013 року по 20 лютого

2018 року, знаходять «... технологічний фактор Hashrate не повинен бути включений в моделювання динаміки цін або фундаментальних значень, оскільки він не впливає на пропозицію біткоїн.

Біткоїн має незалежну «вартість», але не може виміряти зміну цієї вартості, оскільки немає офіційних або неофіційних індексів цін на біткоїн. Крім того, його незначне використання як розрахункової одиниці, засобу обміну та зберігання вартості означає, що біткоїн не задовольняє основні функції, які повинні виконувати "гроші". Тому необхідно дотримуватися обережності при проведенні аналізів у контексті QTM або PPP.

На відміну від його *вартості*, *ціну* біткоїна можна легко виміряти та проаналізувати так само, як і курси іноземних валют. На нинішньому рівні розвитку попит і пропозиція біткоїна дуже нееластичні. Як наслідок, будь-який значний екзогенний шок для попиту чи пропозиції, *ceteris paribus*, викличе нестабільні рухи в ціні, тим самим зменшивши шанси біткоїна діяти як значуще та надійне сховище вартості.

Нарешті, виробничі витрати біткоїна (тобто вартість майнінгу) не мають сильного прямого впливу на сили попиту та пропозиції на валютних ринках біткоїна. У курйозному повороті подій вартість видобутку біткоїни практично не впливає на його ціну. Натомість причинно-наслідковий зв'язок зворотний, зі збільшенням (зменшенням) зростання (зменшення) цін біткоїнів зростанням (зменшенням) витрат майнерів. Цей результат узгоджується з Fantazinni і Kolodin (2020) і Kristoufek (2020), чий економетричні дослідження показують, що «в довгостроковій перспективі витрати на майнінг сильно реагують на зростання ціни біткоїни і в кінцевому підсумку наздоганяють».

Настрої інвесторів виявляються корисними для прогнозування волатильності та прибутковості криптовалюти. Корбет, Ларкін, Люсі, Меган і Яровая (2020) розробили індекс настроїв на основі новинних сюжетів, які слідує за оголошеннями чотирьох макроекономічних показників, тобто ВВП, безробіття, індексу споживчих цін і товарів тривалого користування. Результати показують, що повернення біткоїнів реагує на новини інакше, ніж

прибутковість фондового ринку. Крім того, стверджувалося, що реакція криптовалюти на новини та оголошення може відрізнятися залежно від типу цифрових активів. Тому Корбет, Ларкін і Люсі (2020) стверджували, що цифрові активи на основі валюти виявляються більш сприйнятливими до оголошень про монетарну політику США, тоді як додатки або цифрові активи на основі протоколів несприйнятливі до цих потрясінь. Аналогічна різниця зустрічається для майнінгу і не видобуваються валют, а це означає, що реакція на різні типи невизначеності деяких цифрових активів буде відрізнятися від прибутковості Bitcoin.

Багато літератури з поведінкових фінансів застосували специфічний підхід до вимірювання настроїв інвесторів на ринках криптовалют. Багато досліджень, таких як Лопес-Кабаркос та інші (2019), Уркхарт (2018), Шен, Уркхарт і Ван (2019), а також Ibikunle, McGroarty та Rzaev (2020), Lin (2020) підтвердили, що настрої інвесторів є основним фактором для пояснення та прогнозування прибутковості та волатильності на ринках криптовалют. Задokumentовано двонаправлений причинно-наслідковий зв'язок між увагою інвесторів та поверненням біткоїнів (Dastgir, Demir, Downing, Gozgor, & Lau, 2019). Крім того, використання обсягу пошуку може точно сприяти підвищенню короткострокового прогнозу ринку біткоїнів (Garcia & Schweitzer, 2015; Каралевіцій, Дегранде, & Де Вердт, 2018). Зовсім недавно численні дослідження використовували тенденції Twitter, щоб допомогти виміряти настрої інвесторів. Пізніше багато досліджень показали, що різні аспекти зміни ринку біткоїну, його прибутковостей, волатильності та ліквідності, безсумнівно, можна пояснити на основі Twitter (Shen et al., 2019; Філіппас, Джиба, & Гесмі, 2019; Цой, 2021). Аналогічні результати також були отримані для високочастотних даних, оскільки статистично значущі зв'язки між настроями інвесторів та прибутковістю Bitcoin визначені Gu'egan та Renault (2021). Інші відносини між увагою інвесторів та виявленням цін помічені на криптовалютних ринках після вивчення Ibikunle та ін. (2020 р.). Крім того, крім Twitter і Google Trends, дослідники з академічної області також

виявили нові можливі змінні у вимірюванні уваги інвесторів. Наприклад, можна використовувати кілька нових майданчиків, щоб привернути увагу інвесторів, і ця нова змінна надзвичайно важлива для прибутковості та волатильності (Sabah, 2020).

#### **1.4 Висновки до розділу 1**

Криптоекономіка характеризується істотною відмінністю від інших фінансових ринків. Вона включає криптовалюти, які є виключно мінливими з майже однаковою ймовірністю величезної прибутковості або з однаковою ймовірністю величезних втрат. Результати показують надзвичайно високу мінливість цін на криптовалюти, і ця мінливість значною мірою не пов'язана з коливаннями цін на інших стандартних фінансових ринках. У будь-якому випадку, винахід криптовалют, особливо використання технології блокчейн, змінює наше сприйняття традиційних грошей. Крім того, криптовалюти сприяли інноваціям і проникненню в невідомі області фінансових ринків. Вони дозволили величезній кількості людей брати участь у цьому ринку, розпочати торгівлю та інвестувати свої гроші з мінімальними зусиллями чи знаннями.

Тим не менш, оскільки ринок криптовалют не регулюється жодним законодавчим органом, він піддатливий до різноманітних «пампів» та «дампів», що ставить вкладення багатьох інвесторів під загрозу. Тому постає актуальність у впровадженні підходів на основі теорії складності для передчасної ідентифікації кризових подій.

## РОЗДІЛ 2. МОЖЛИВІСТЬ ВИКОРИСТАННЯ ІНСТРУМЕНТАРІЮ ЕНТРОПІЙНИХ ПОКАЗНИКІВ СКЛАДНОСТІ

### 2.1 Ентропія. Апроксимаційна ентропія та ентропія шаблонів

Ентропія є мірою неорганізованості та непередбачуваності складної системи. Чим більшу ентропію система має, тим менш передбачуваною вона представляється. Можна думати про ентропію як про випадковість (хоча справжня випадковість дуже рідкісна, тому на практиці ми часто задовольняємося псевдовипадковістю).

Метою апроксимаційної ентропії (ApEn) та ентропії шаблонів (SampEn) є оцінка випадковості ряду даних без будь-яких попередніх знань про джерело, що генерує набір даних. Тому їх застосовність безмежна, і ці алгоритми використовувалися в найрізноманітніших дослідницьких областях. На жаль, через простоту алгоритмів іноді теорія, що стоїть за ними, невідома дослідникам, обмежуючи можливості подальшого вдосконалення галузі або кращого розуміння що ці методики вимірюють.

Щоб вивчити поняття випадковості ряду, Чайтін запропонував експеримент з перевертання монети 20 разів і записати один, якщо він прийде головами, і нуль, якщо справа дійде хвостами. З імовірнісної точки зору експеримент може призвести до наступних двох серій з однаковою ймовірністю (один над  $2^{20}$ ), таким чином розглядаючи обидва ряди як випадкові (зверніть увагу, що оригінальна серія В на прикладі Чайтіна є 01101100110111100010; Однак цей ряд має інше середнє значення, ніж серія А. Щоб підкреслити той факт, що два ряди не відрізняються з точки зору середнього значення, дисперсії тощо, ми змінили серію В, щоб мати десять одиниць і десять нулів, як серія А):

В: 010101010101010101

Б: 01101000110111100010

Хоча це правда, що два ряди були створені випадковим чином з однаковою ймовірністю і мають однакові значення середнього та дисперсійного значення, легко продовжити закономірність, описану в серії А, але нам доведеться попрацювати більше, щоб мати можливість передбачити наступне число в серії В.

Серія А будується шляхом чергування нулів і одиниць, в той час як серія В не має чіткого малюнка. Таким чином, аналіз випадковості є істотним обмеженням методів статистики моментів і не може бути використаний для гарантування випадковості ряду математично.

По суті, оскільки ентропія кількісно визначає обсяг інформації, вона також вимірює ступінь випадковості в системі. Хотілося б, щоб однією з характеристик при вимірюванні випадковості ряду було встановлення ієрархії ступенів випадковості, кількісного визначення рівня випадковості так, щоб питання полягало не в тому, чи є ряд даних випадковим чи ні, а в тому, наскільки він випадковий. Буде ситуація з максимальною випадковістю, абсолютно непередбачуваною, але також будуть інші ситуації з різним ступенем випадковості, тому ми повинні мати можливість упорядкувати ряд даних відповідно до їх складності. Простий (не складний) ряд даних мав би очевидні закономірності, такі як вищезгаданий ряд А, тоді як для складного ряду ці закономірності були б абсолютно відсутніми.

Складність ряду даних можна оцінити за допомогою таких алгоритмів, як  $ApEn$  або  $SampEn$ . Теоретичні ідеї, що стоять за ними, мають коріння в теорії інформації та концепції ентропії, величина якої дозволяє вимірювати кількість інформації математичним способом. Використовуючи ідеї ІТ, дослідники теорії хаосу сформулювали різні способи визначення ентропії з наміром класифікувати хаотичні системи. Однак, незважаючи на те, що ці рівняння корисні для детермінованих процесів, невеликі кількості шуму роблять їх недійсними для реальних рядів даних.

Хоча представлені алгоритми спочатку розроблялися для клінічних досліджень, пізніше вони поширилися на різноманітні галузі, такі як

нейроінженерія, візуальне розпізнавання образів, нейроінформатика, екологія, психіатрія, електроніка, розпізнавання голосу або фінанси. Оскільки останнім часом їх застосування зростає, у цій роботі ми представили чіткий шлях для розуміння логіки  $ApEn$  та  $SampEn$ , щоб допомогти дослідникам зрозуміти їх основи та правильне застосування.

У багатьох ситуаціях дослідники використовують представлені алгоритми, не знаючи основ, що стоять за ними, що може призвести до неправильної інтерпретації результатів. Наприклад, поширена помилка полягає в тому, що  $ApEn$  забезпечує більшу регулярність, ніж реальне значення. Однак це твердження вірне лише в тому випадку, якщо вибране значення  $r$  знаходиться перед максимумом  $ApEn$ ; за межами цієї точки  $SampEn$  забезпечує більшу регулярність, ніж дійсне значення, завдяки своєму формулюванню на основі ентропії Ренні порядку два.

$ApEn$  є мірою регулярності, що визначає можливість передбачувати флуктуації в часових рядах. Інтуїтивно вона означає, що наявність повторюваних шаблонів (послідовностей певної довжини, побудованих із чисел ряду, що слідує одне за іншим) флуктуацій у часовому ряді призводить до більшої передбачуваності часового ряду порівняно із рядами, де повторюваності шаблонів немає. Порівняно велике значення  $ApEn$  показує ймовірність того, що подібні між собою шаблони спостережень не будуть слідувати один за одним. Іншими словами, часовий ряд, що містить велику кількість повторюваних шаблонів, має порівняно мале значення  $ApEn$ , а значення  $ApEn$  для менш передбачуваного (більш складного) процесу є більшим.

При розрахунку  $ApEn$  для даного часового ряду  $S_N$ , що складається з  $N$  значень  $t(1), t(2), t(3), \dots, t(N)$  вибираються два параметри,  $m$  та  $r$ . Перший з цих параметрів –  $m$  – вказує на довжину шаблону, а другий –  $r$  – визначає критерій подібності. Досліджуються підпослідовності елементів часового ряду  $S_N$ , що складаються з  $m$  чисел, взятих, починаючи з номера  $i$  та

називаються векторами  $p_m(i)$ . Два вектори (шаблони),  $p_m(i)$  та  $p_m(j)$  будуть подібними, якщо всі різниці пар їх відповідних координат є меншими за значення  $r$ , тобто якщо

$$|t(i+k) - t(j+k)| < r \text{ для } 0 \leq k < m.$$

Для розглядуваної множини  $P_m$  всіх векторів довжини  $m$  часового ряду  $S_N$  можна обраховуються значення

$$C_{im}(r) = \frac{n_{im}(r)}{N - m + 1},$$

де  $n_{im}(r)$  – кількість векторів у  $P_m$ , що подібні вектору  $p_m(i)$  (враховуючи вибраний критерій подібності  $r$ ). Значення  $C_{im}(r)$  є часткою векторів довжини  $m$  що мають схожість із вектором такої ж довжини, елементи якого починаються з номера  $i$ . Для даного часового ряду обраховуються значення  $C_{im}(r)$  для кожного вектора у  $P_m$ , після чого знаходиться середнє значення  $C_m(r)$ , яке виражає розповсюдженість подібних векторів довжини  $m$  у ряду  $S_N$ . Безпосередньо ентропія подібності для часового ряду  $S_N$  з використанням векторів довжини  $m$  та критерію подібності  $r$  визначається за формулою:

$$ApEn(S_N, m, r) = \ln \left( \frac{C_m(r)}{C_{m+1}(r)} \right),$$

тобто, як натуральний логарифм відношення повторюваності векторів довжиною  $m$  до повторюваності векторів довжиною  $m+1$ .

Таким чином, якщо знайдуться подібні вектори у часовому ряді,  $ApEn$  оцінить логарифмічну ймовірність того, що наступні інтервали після кожного із векторів будуть відрізнятись. Менші значення  $ApEn$  відповідають більшій ймовірності того, що за векторами слідує подібні їм. Якщо часовий ряд дуже нерегулярний – наявність подібних векторів не може бути передбачуваною і значення  $ApEn$  є порівняно великим.



Зауважимо, що  $ApEn$  є нестійкою до вхідних даних характеристикою, оскільки досить сильно залежить від параметрів  $m$  та  $r$ .

## 2.2 Пермутаційна ентропія: природна міра складності часових рядів

Були розроблені різні міри складності для порівняння часових рядів і розрізнення регулярної (наприклад, періодичної), хаотичної і випадкової поведінки. Серед іншого, повідомлялося, що складність даних серця і мозку може розрізняти здорових і хворих суб'єктів, а іноді навіть прогнозувати інфаркт або епілептичний випадок [1]. Основними типами параметрів складності є ентропії, фрактальні розміри і показники Ляпунова. Всі вони визначені для типових орбіт імовірно ергодичних динамічних систем, і між цими величинами існують глибокі зв'язки.

Основна концептуальна проблема полягає в тому, що ці визначення не зроблені для довільного ряду спостережень  $x_1, x_2, \dots$ . Як наслідок, існує і обчислювальна задача. Протягом останніх 20 років було розроблено багато геніальних алгоритмів, хитрощів і рецептів для оцінки показників складності з реальних часових рядів.

Вони чудово працюють, коли часові ряди моделюються з низьковимірної динамічної системи, але більшість з них ламається, як тільки в серію додається шум. Для реальних серій «усунення шуму» вимагає ретельної попередньої обробки даних і доопрацювання параметрів, а результати неможливо відтворити без вказівки деталей методу.

Наші ентропії обчислюються для різної розмірності вкладень  $n$ , але ми не намагаємося визначити межу для великих  $n$ , хоча це цікава теоретична задача [16]. Для практичних цілей рекомендуємо  $n = 3, \dots, 7$ .

Приклад: Візьмемо ряд з сімома значеннями:

$$x = (4, 7, 9, 10, 6, 11, 3).$$

Організуємо шість пар сусідів, відповідно до їх відносних значень, знаходячи чотири пари, для яких  $x_t < x_{t+1}$  і дві пари, для яких  $x_t > x_{t+1}$ . Так чотири з шести пар значень представлені перестановкою 01 ( $x_t < x_{t+1}$ ) і дві з шести представлені 10. Визначимо перестановочну ентропію порядку  $n = 2$  як міру ймовірностей перестановок 01 і 10. Так

$$H(2) = -(4/6) \log(4/6) - (2/6) \log(2/6) \approx 0.918.$$

Як завжди, логарифм з основою 2, тому  $H$  визначається в бітах. Далі порівнюємо три послідовних значення. (4, 7, 9) і (7, 9, 10) представляють перестановку 012, оскільки вони знаходяться в зростаючому порядку. (9, 10, 6) і (6, 11, 3) відповідають перестановці 201, оскільки  $x_{t+2} < x_t < x_{t+1}$ , тоді як (10, 6, 11) відповідатиме перестановці 102 ( $x_{t+1} < x_t < x_{t+2}$ ). Пермутаційна ентропія порядку  $n = 3$  дорівнює  $H(3) = -2(2/5) \log(2/5) - (1/5) \log(1/5) \approx 1.522$ .

Виходячи з вище зазначено, алгоритм визначення пермутаційної ентропії можна представити наступним чином. Розглянемо часовий ряд  $\{x_t\}_{t=1, \dots, T}$ . Вивчимо всі  $n!$  перестановок  $\pi$  порядку  $n$ . Для кожного  $\pi$  визначаємо відносну частоту (# означає число)

$$p(\pi) = \frac{\#\{t | t \leq T - n, (x_{t+1}, \dots, x_{t+n}) \text{ has type } \pi\}}{T - n + 1}.$$

Для визначення  $p(\pi)$  ми маємо врахувати випадок, коли  $T \rightarrow \infty$  у наведеній вище формулі. Ця межа існує з імовірністю 1, коли основний стохастичний процес виконує дуже слабку умову стаціонарності: для  $k \leq n$ , ймовірність  $x_t < x_{t+k}$  не повинна залежати від  $t$ .

Пермутаційна ентропія порядку  $n \geq 2$  визначається як

$$H(n) = - \sum p(\pi) \log p(\pi),$$

де сума проходить над усіма  $n!$  перестановками  $\pi$  порядку  $n$ . Це інформація, що міститься при порівнянні  $n$  послідовних значень часового ряду. Зрозуміло, що  $0 \leq H(n) \leq \log n!$ , де нижня межа досягається для

зростаючої або спадної послідовності значень, а верхня межа для повністю випадкової системи, де всі  $n!$  можливих перестановок мають однакову ймовірність виникнення. Часовий ряд представляє якусь динаміку, коли  $H(n) < \log n!$ .

**Висновок.** Перестановочна ентропія є відповідним заходом складності для хаотичних часових рядів, зокрема за наявності динамічного та спостережного шуму. На відміну від всіх відомих параметрів складності, невеликий шум істотно не змінює складності хаотичного сигналу. Перестановочні ентропії можна обчислити для довільних реальних часових рядів. Оскільки метод надзвичайно швидкий і надійний, він здається кращим, коли є величезні набори даних і немає часу на попередню обробку і тонку настройку параметрів.

### 2.3 Дисперсійна ентропія. Міра для аналізу часових рядів

Одним з найпотужніших інструментів оцінки динамічних характеристик часових рядів є ентропія. Ентропія шаблонів хоч і потужна, але є недостатньо швидкою, особливо для довгих сигналів. Пермутаційна ентропія, як широко використовуваний показник, розглядає тільки порядок значень і тому деяка інформація щодо амплітуд може бути відкинута.

Для вирішення цих проблем ми вводимо новий метод, який називається дисперсійною ентропією (DE), для кількісної оцінки регулярності часових рядів. Ми отримуємо уявлення про залежність DE від декількох простих концепцій обробки сигналів за допомогою набору синтетичних часових рядів. Результати показують, що DE, на відміну від PE, може виявляти смугу пропускання шуму і одночасну зміну частоти і амплітуди.

**Дисперсійна ентропія.** Для заданого одновимірного сигналу довжини  $N$ :  $x = \{x_1, x_2, \dots, x_N\}$ , алгоритм DE включає в себе 4 основних кроки:

1) По-перше,  $x_j (j = 1, 2, \dots, N)$  зіставляються з класами  $c$ , позначеними від 1 до  $c$ . Для цього існує ряд лінійних і нелінійних підходів. Хоча алгоритм лінійного відображення є найшвидшим, коли максимальні та/або мінімальні значення часового ряду набагато більші або менші за середнє/медіанне значення сигналу, більшості  $x_i$  присвоюється лише кілька класам.

Таким чином, ми вперше використовуємо функцію нормального інтегрального розподілу (NCDF) для відображення  $x$  в  $y = \{y_1, y_2, \dots, y_N\}$  від 0 до 1. Далі використовуємо лінійний алгоритм для присвоєння кожному  $y_i$  цілого числа від 1 до  $c$ . Для цього до кожного члена відображеного сигналу ми використовуємо  $z_j^c = \text{round}(c \cdot y_j + 0.5)$ , де  $z_j^c$  представляє  $j^{\text{th}}$  член класифікованого часового ряду, а округлення передбачає або збільшення, або зменшення числа до наступної цифри. Варто відзначити, що цей крок міг бути здійснений і деякими іншими лінійними і нелінійними методами відображення.

2) Кожен вектор вкладень  $z_i^{m,c}$  з розмірністю вкладень  $m$  та часовою затримкою  $d$  відповідає  $z_i^{m,c} = \{z_i^c, z_{i+d}^c, \dots, z_{i+(m-1)d}^c\}$ ,  $i = 1, 2, \dots, N - (m - 1)d$ . Кожен вектор  $z_i^{m,c}$  проектується на дисперсійний шаблон  $\pi_{v_0 v_1 \dots v_{m-1}}$ , де  $z_i^c = v_0$ ,  $v_{i+d}^c = v_1, \dots, z_{i+(m-1)d}^c = v_{m-1}$ . Кількість можливих дисперсійних шаблонів, що можуть бути присвоєнні до кожного вектора  $z_i^{m,c}$  дорівнює  $c^m$ , оскільки сигнал має  $m$  значень і кожне значення може приймати клас 1 до  $c$ .

3) Відносну частоту кожного потенційного дисперсійного шаблону можна визначити як

$$p(\pi_{v_0 v_1 \dots v_{m-1}}) = \frac{\#\{i | i \leq N - (m - 1)d, z_i^{m,c} \text{ has type } \pi_{v_0 v_1 \dots v_{m-1}}\}}{N - (m - 1)d}$$

4) Нарешті, опираючись на класичну Шенонівську ентропію, значення DE з розмірністю вкладень  $m$ , часовою затримкою  $d$  та кількістю класів  $c$  можна визначити наступним чином:

$$DE(x, m, c, d) = - \sum_{\pi=1}^{c^m} p(\pi_{v_0 v_1 \dots v_{m-1}}) \ln p(\pi_{v_0 v_1 \dots v_{m-1}}).$$

Коли всі дисперсійні шаблони розподілені рівномірно, DE приймає найбільше значення -  $\ln c^m$ . На противагу до цього, якщо у розподілі  $\pi_{v_0 v_1 \dots v_{m-1}}$  присутнє лише одне значення відмінне від нуля, тоді досліджуваний сигнал представляється достатньо регулярним/передбачуваним, а значення DE є найменшим.

Дисперсійні закономірності в DE відповідають закономірностям перестановки в PE, хоча дисперсійні шаблони, на відміну від перестановки, включають рівні амплітудні значення і в векторах вкладень. Коли значення ряду є рівними, Бандт і Помбе [2] запропонували ранжувати можливі рівності на основі порядку їх виникнення або вирішення цієї умови шляхом додавання шуму. З огляду на першу альтернативу, наприклад, шаблон перестановки як для  $\{1, 2, 4\}$ , так і для  $\{1, 4, 4\} \in 012$ . Припустимо  $q_1 = \{1, 2, 2, 2\}$  і  $q_2 = \{1, 2, 3, 4\}$ . Схема перестановки  $q_1$  точно така ж, як і у  $q_2$ , хоча, на відміну від  $q_1$ ,  $q_2$  строго висхідна. Додавання шуму може не призвести до правильної корекції, оскільки, наприклад,  $\{1, 4, 4\}$  має дві можливі моделі перестановки як 012 та 021, і між ними немає жодних відмінностей. Ця проблема особливо актуальна для оцифрованих сигналів. Слід додати, що ця проблема існує і в зваженій пермутаційній ентропії, яка була запропонована нещодавно [9]. На відміну від цього, в методах DE з  $c=3$  і  $m=3$  ми маємо потенційні закономірності на кшталт 122 і 222.

Важливим питанням у кожному підході, заснованому на ентропії, є вибір відповідних значень параметрів. Тут ми маємо три параметри, включаючи розмірність вкладень  $m$ , часову затримку  $d$  та кількість класів  $c$ . Для практичних цілей DE, як і PE, ми рекомендуємо припускати  $d = 1$ , оскільки при  $d > 1$  деяка важлива інформація з точки зору частоти може бути відкинута. Зрозуміло, що  $c$  має бути більшим за 1, тому що при  $c = 1$  існує лише одна схема дисперсії. Для роботи з достовірною статистикою при

розрахунку DE пропонується, щоб кількість потенційних класів дисперсії ( $c^m$ ) була менше за довжину сигналу. Коли  $c$  занадто мале, два значення амплітуди, які знаходяться далеко один від одного, можуть бути віднесені до подібного класу, в той час як занадто велике значення  $c$ , дуже невеликі відмінності можуть змінити їх клас, і тому метод DE може бути чутливим до шуму. Якщо шуму немає, вибрати великий  $c$  має сенс. Більше того, коли  $m$  або  $c$  занадто великі, час обчислень високий, хоча це робить значення DE більш надійними. Крім того, якщо розмірність вкладень  $m$  занадто мала, динамічні зміни можуть бути і не виявлені в сигналі, тоді як великі  $m$  можуть призвести до того, що алгоритм DE не може врахувати невеликі варіації.

#### **2.4 Нечітка ентропія. Вимірювання регулярності часових рядів за допомогою нелінійної вибірки на основі подібності ентропії.**

Найбільш прямим зв'язком між теорією хаосу і реальним світом є аналіз часових рядів з реальних систем з точки зору нелінійної динаміки. Пінкус запропонував показник названий апроксимаційною ентропією, спрямоване на вимірювання регулярності сигналу, тобто вимірювання наявності схожих закономірностей в часовому ряду. Як помітив сам Пінкус, на апроксимаційну ентропію впливає ефект зміщення.

Річман і Моорман розробили модифікацію алгоритму ApEn, названого ентропією шаблонів (SampEn), щоб усунути деякі недоліки, про які повідомлялося вище. SampEn показує відносну узгодженість у випадках, коли ApEn цього не робить.

Він потенційно може розрізнити найрізноманітніші системи: низьковимірні детерміновані системи, періодичні, високовимірні хаотичні системи, а також стохастичні системи. Даний показник широко застосовувався до різноманітних фізіологічних і клінічних даних часових

рядів, таких як сигнали серцевого ритму, сигнали артеріального тиску, електрокардіограма, електроенцефалографія, електроміографія тощо.

Тим не менш, значення SampEn є переривчастим і, можуть значно змінюватись при незначній зміні порогу  $r$  для функції Гевісайда, що використовується для визначення подібності векторів. У даному підрозділі представляється опис модифікована ентропії шаблонів на основі нелінійної сигмоїдної функції, яка визначає подібність двох векторів

Для  $N$  значень ряду  $\{u(i): 1 \leq i \leq N\}$  та розмірності вкладень  $m$ , ми формуємо вектори вкладень  $\{X_i^m, i = 1, \dots, N - m + 1\}$  наступним чином:

$$X_i^m = \{u(i), u(i + 1), \dots, u(i + m - 1)\} - u_0(i),$$

де  $X_i^m$  представляє  $m$  послідовних незміщених значень, де

$$u_0(i) = \frac{1}{m} \sum_{j=0}^{m-1} u(i + j).$$

Для певного вектора  $X_j^m$  потрібно визначити відстань  $d_{ij}^m$  між  $X_i^m$  та  $X_j^m$  як максимальну абсолютну відстань між відповідними скалярними компонентами векторів:

$$d_{ij}^m = d[X_i^m, X_j^m] = \max_{k \in (0, m-1)} |(u(i + k) - u_0(i)) - (u(j + k) - u_0(j))|.$$

Для значень  $n$  та  $r$  потрібно розрахувати ступінь подібності  $D_{ij}^m$  між  $X_i^m$  та  $X_j^m$  через нечітку функцію  $\mu(d_{ij}^m, n, r)$ :

$$D_{ij}^m(n, r) = \mu(d_{ij}^m, n, r),$$

де нечітка функція  $\mu(d_{ij}^m, n, r)$  представляється експоненційною функцією:

$$\mu(d_{ij}^m, n, r) = \exp(-(d_{ij}^m)^n / r).$$

Далі потрібно визначити функцію  $\phi^m$  як

$$\phi^m(n, r) = \frac{1}{N - m} \sum_{i=1}^{N-m} \left( \frac{1}{N - m - 1} \sum_{j=1, j \neq i}^{N-m} D_{ij}^m \right).$$

Подібним чином ми формуємо  $\phi^{m+1}$ :

$$\phi^{m+1}(n, r) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left( \frac{1}{N-m-1} \sum_{j=1, j \neq i}^{N-m} D_{ij}^{m+1} \right).$$

У підсумку, ми можемо визначити нечітку ентропію (FuzzEn) досліджуваної послідовності як від'ємний натуральний логарифм відхилення  $\phi^m$  від  $\phi^{m+1}$ :

$$\text{FuzzEn}(m, n, r, N) = \ln \phi^m(n, r) - \ln \phi^{m+1}(n, r).$$

## 2.5 Висновки до розділу 2.

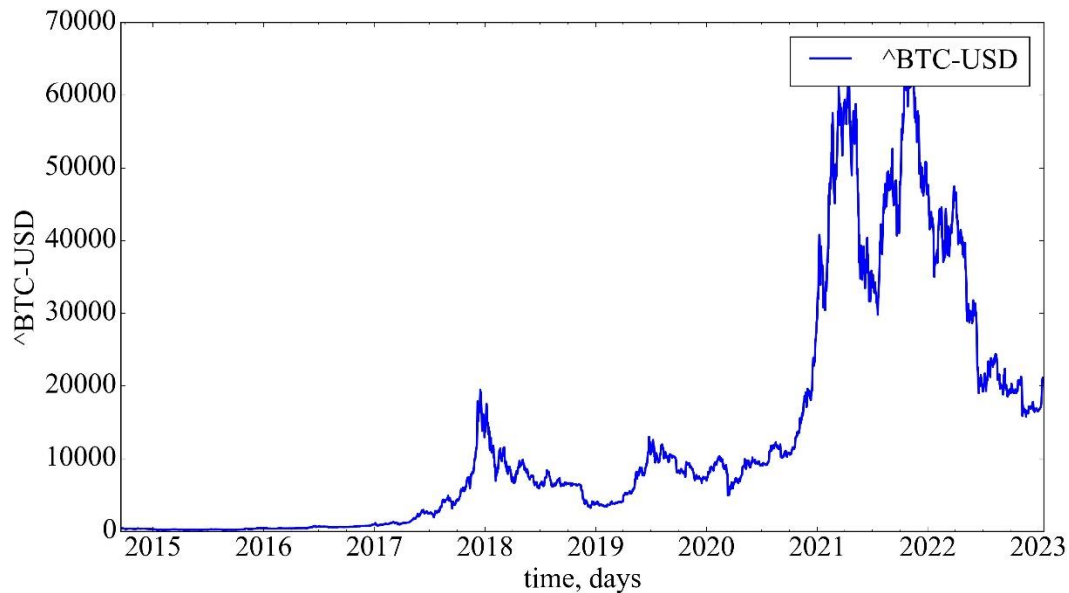
Отже, Функція Sigmoid виступає в ролі вихідних воріт, які можуть бути відкриті або закриті. Оскільки функція неперервна, також можна частково відкрити ворота (тобто десь між 0 і 1). Використовуючи функцію Sigmoid, нелінійна подібність, заснована на mSampEn, змінює правило визначення подібності векторів в SampEn, яке базується на функції Heaviside. На відміну від жорсткої і переривчастої межі функції Хевісайда, м'яка і неперервна границя сигмовидної функції змушує статистику mSampEn плавно змінюватися при незначній зміні допуску  $r$ .

Згідно з визначенням, mSampEn перевершує SampEn за відносною узгодженістю, свободою вибору параметрів, стійкістю до шуму та незалежністю від довжини даних при характеристиці часових рядів з різними закономірностями. В результаті статистика mSampEn забезпечує поліпшену оцінку регулярності часових рядів, і, таким чином, може бути більш зручно і потужно застосована до галасливих фізіологічних сигналів з короткою довжиною даних.

## РОЗДІЛ 3. ПОРІВНЯЛЬНИЙ АНАЛІЗ ПРЕДСТАВЛЕНИХ ПОКАЗНИКІВ, ОПИРАЮЧИСЬ НА ЧАСОВИЙ РЯД БІТКОЇНА.



У даному розділі ми досліджуватимемо динаміку інформаційних показників складності для часового ряду біткоїна. Ряд було завантажено з сайту Yahoo Finance за весь часовий період (Рис. 3.1).



**Рис. 3.1.** Часовий ряд біткоїна за весь період, що представляє сайт Yahoo Finance.

Кожен із представлених показників розраховуватиметься з використанням алгоритму ковзного (рухомого) вікна. Суть алгоритму полягає у тому, що береться фрагмент (вікно) ряду певної довжини. Для даного фрагменту розраховується міра складності й далі вікно зміщується на певний передвизначений крок. Розрахунки повторюються до повного вичерпання ряду. Всі збережені значення отримані з процедури рухомого вікна співставляються з динамікою досліджуваного ряду. Опираючись на динаміку розрахованого показника, ми можемо робити висновки стосовно теперішнього стану системи та її подальшого руху.

Розрахунки даних показників виконувались із використанням процедури грануляції. Ідея класичного підходу включає процедури «грубого дроблення» (coarse graining – «грануляції») початкового часового ряду – усереднення даних на сегментах, що не перетинаються, розмір яких (вікно

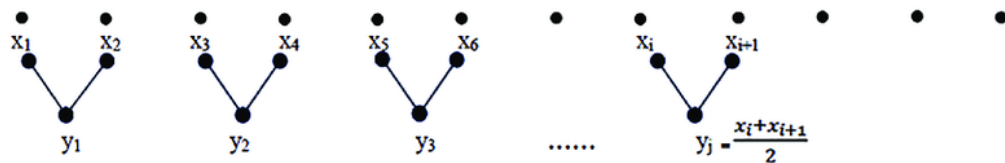
усереднення) збільшувався на одиницю при переході на наступний за величиною масштаб та обчислення на кожному з масштабів певного ентропійного показника.

Кожен елемент «гранульованого» часового ряду знаходиться у відповідності до виразу:

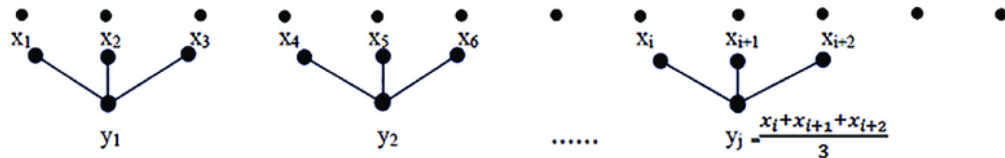
$$y_j^\tau = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i, \quad 1 \leq j \leq N/\tau,$$

де  $\tau$  характеризує показник масштабування. Довжина кожного «гранульованого» ряду залежить від розміру вікна і рівна  $N/\tau$ . Для  $\tau = 1$  «гранульований» ряд тотожний оригінальному. На рис. 3.2 представлена орієнтовна схема процедури грануляції.

### Scale 2



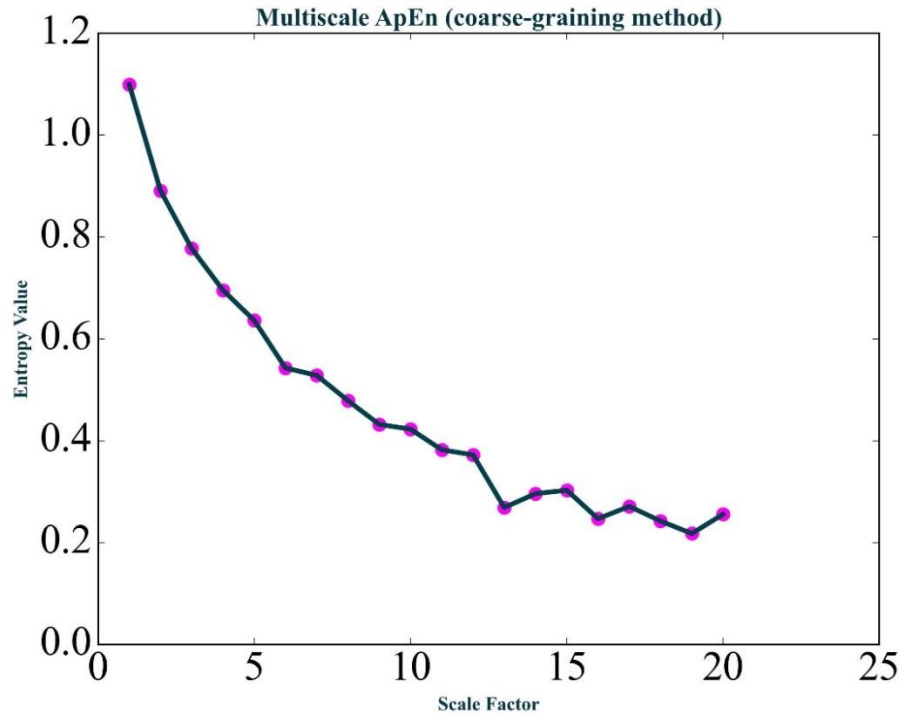
### Scale 3



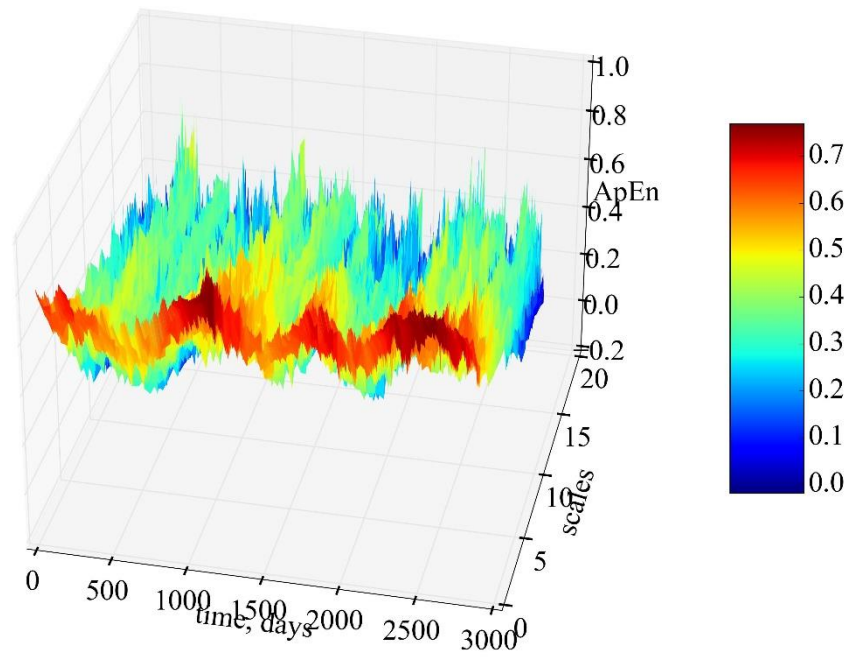
**Рис. 3.2.** Схематична ілюстрація процедури грануляції часового ряду для масштабів 2 і 3.

## 3.1 Апроксимаційна ентропія

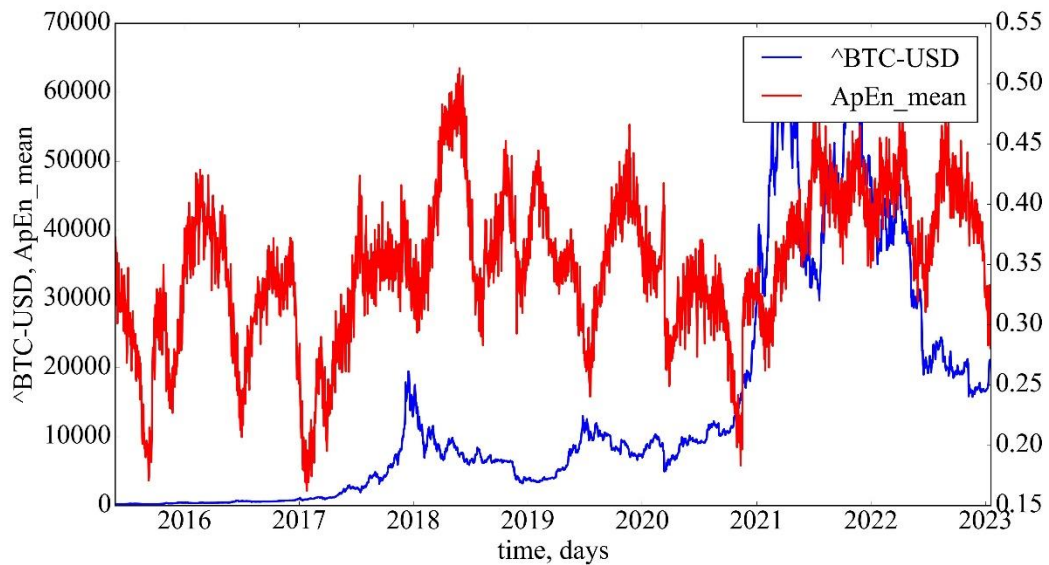
Продемонструємо, якою представлятиметься складність біткоїна для різних часових масштабів. На рис. 3.3 представлено мультимасштабну апроксимаційну ентропію для всього часового ряду, для перших 20 масштабів, та з врахуванням алгоритму рухомого вікна.



(a)



(b)



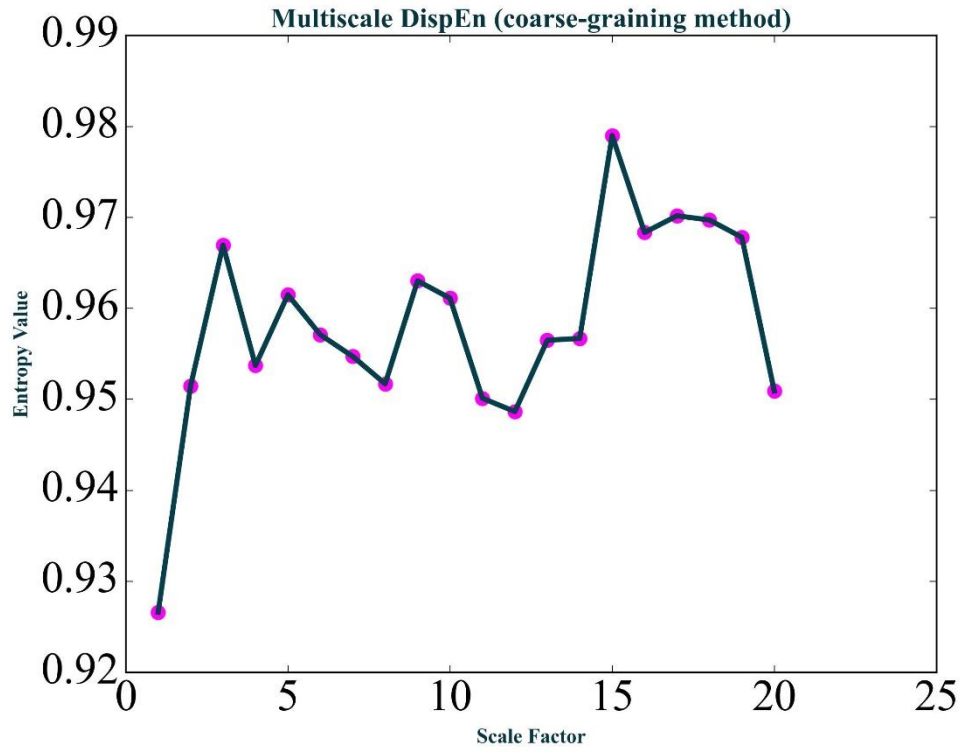
(в)

**Рис. 3.4.** Залежність апроксимаційної ентропії (Entropy Value) від масштабу (Scale Factor) (а), мультискейлінгова поверхня апроксимаційної ентропії (б) та усереднена по всім масштабам апроксимаційна ентропія (в).

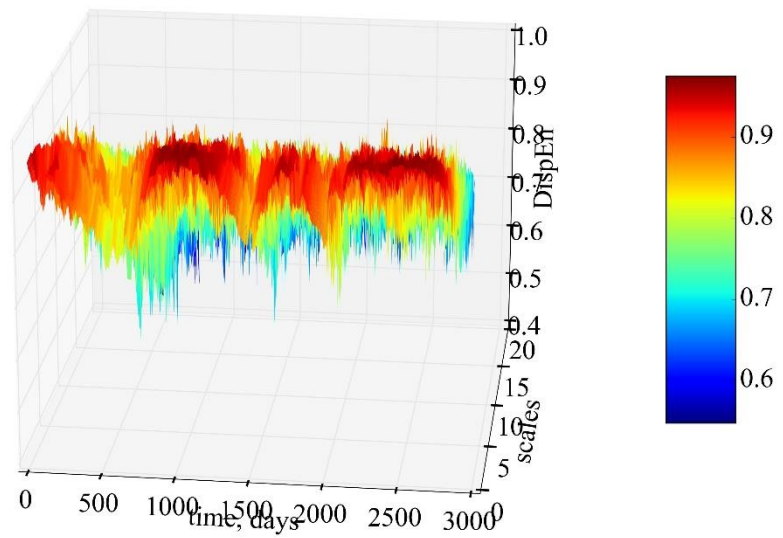
Видно, що з процедурою грануляції складність біткоїна спадає. Усереднення з використанням більших масштабів говорить про те, що біткоїн представляється більш передбачуваним. Навіть із грануляцією ряду, ми маємо змогу спостерігати зростання апроксимаційної ентропії під час криз або в передкризовий період. Це може слугувати індикатором або індикатором-передвісником.

### 3.2 Дисперсійна ентропія

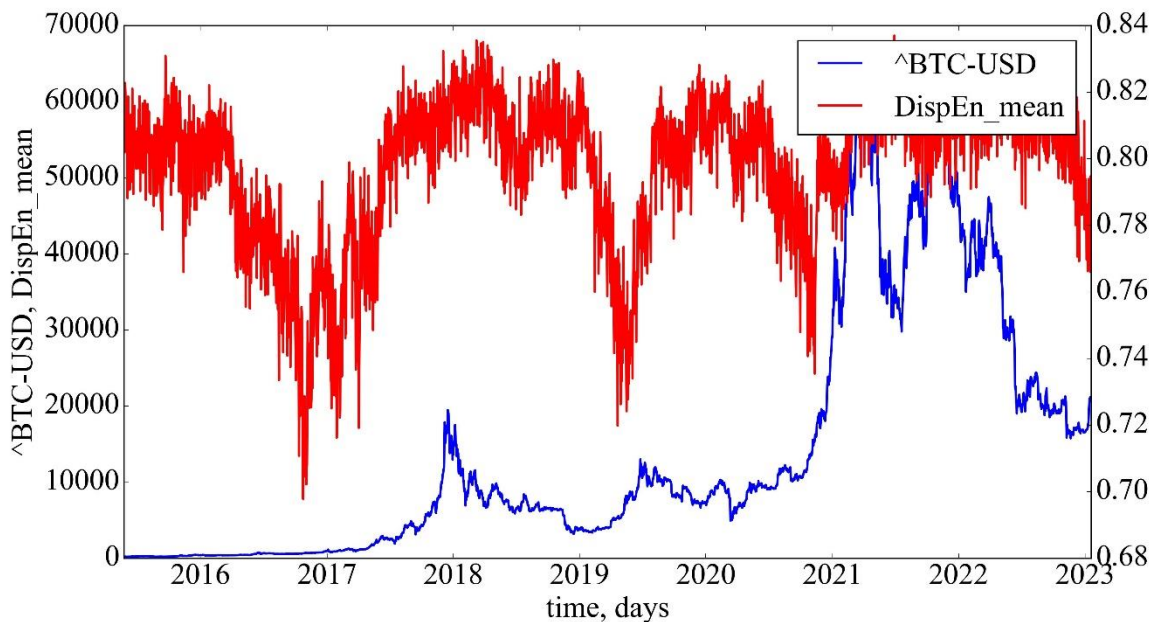
На рис. 3.4 представлено мультимасштабну дисперсійну ентропію для всього часового ряду, з врахуванням перших 20 масштабів та на основі алгоритму рухомого вікна.



(a)



(b)

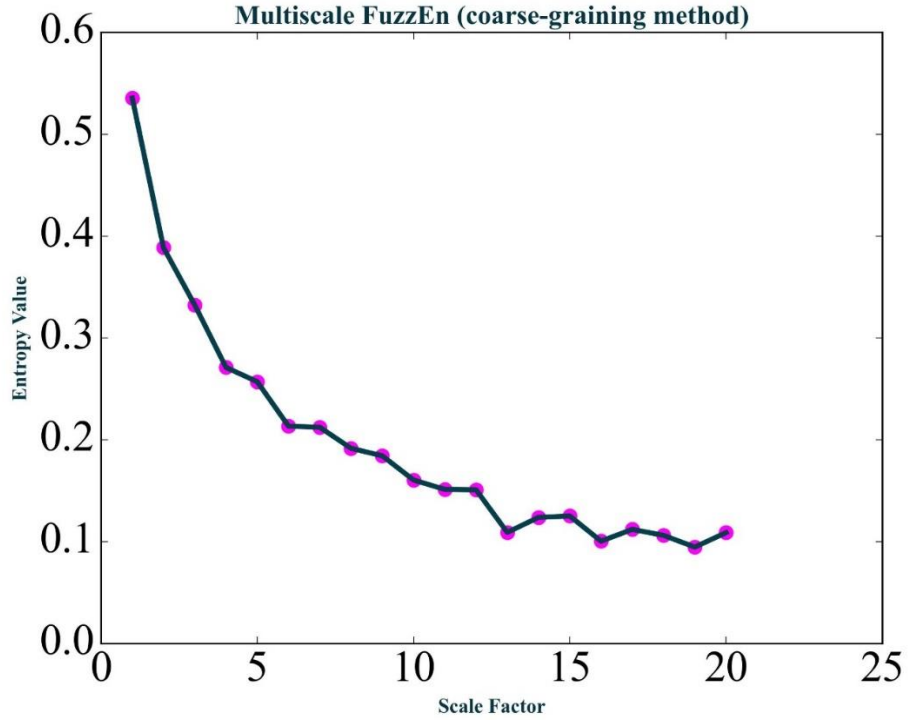


(в)

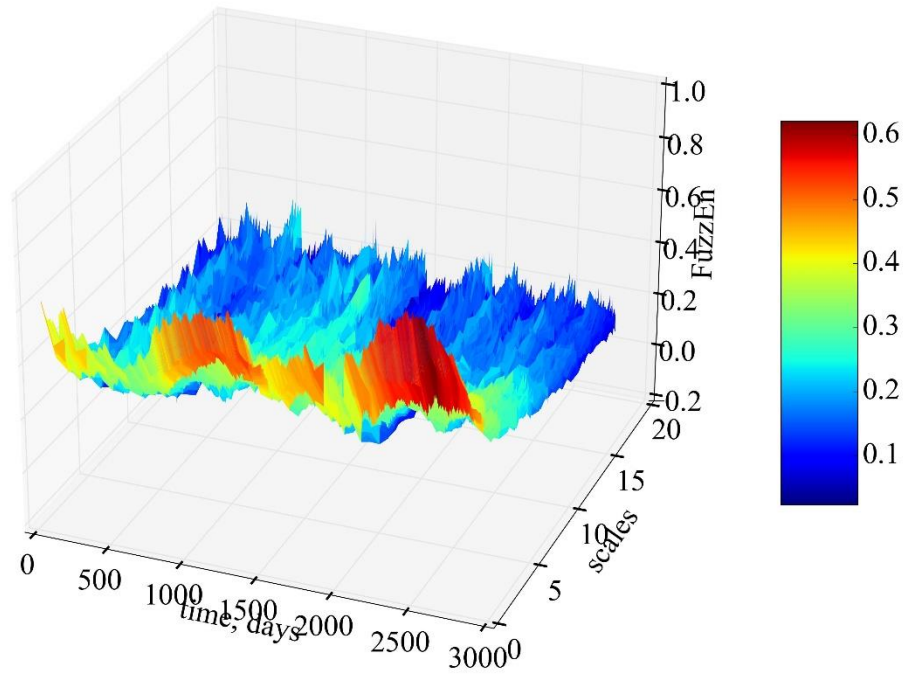
**Рис. 3.4.** Залежність дисперсійної ентропії (Entropy Value) від масштабу (Scale Factor) (а), мультискейлінгова поверхня дисперсійної ентропії (б) та усереднена по всім масштабам дисперсійна ентропія (в).

### 3.3 Нечітка ентропія

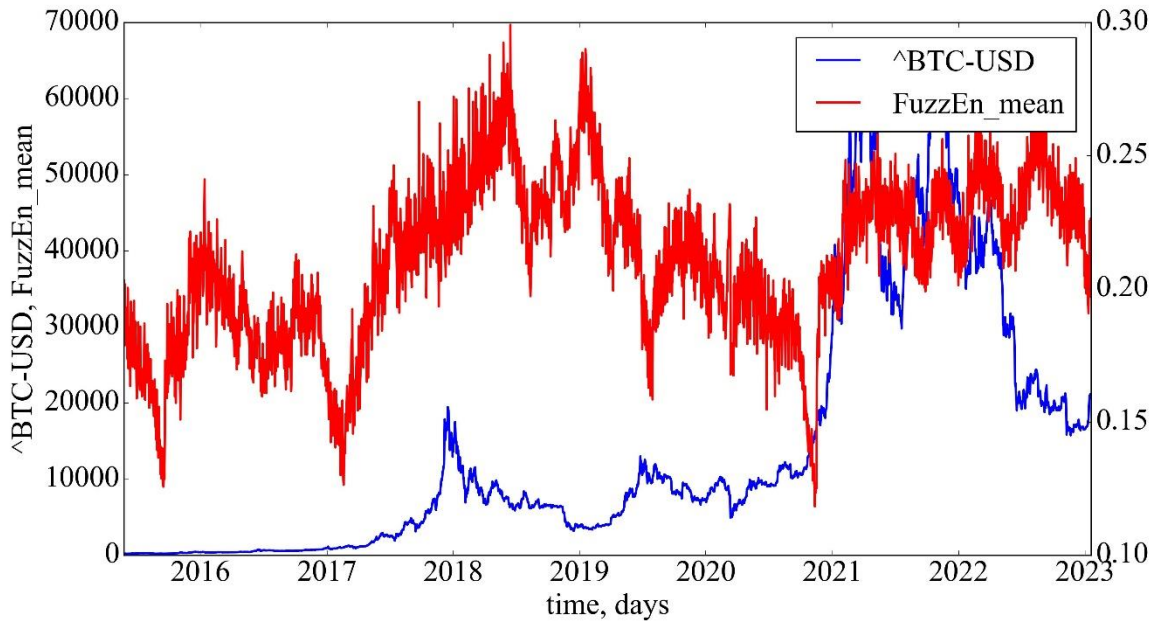
На рис. 3.5 представлено мультимасштабну нечітку ентропію для всього часового ряду, з врахуванням перших 20 масштабів та на основі алгоритму рухомого вікна.



(a)



(b)



(в)

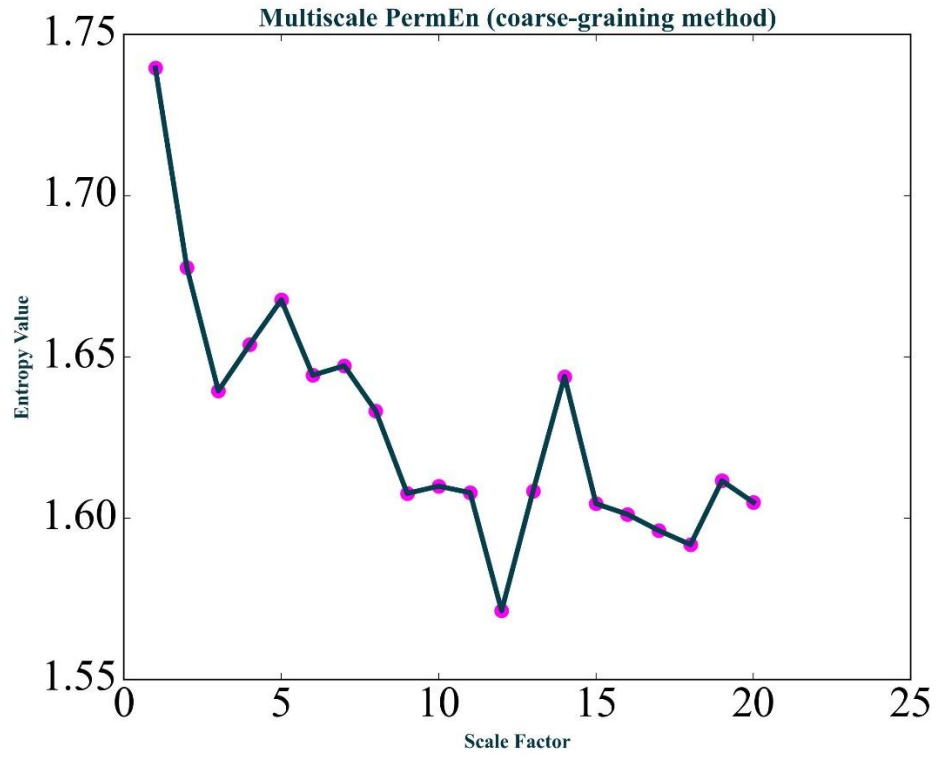
**Рис. 3.5.** Залежність нечіткої ентропії (Entropy Value) від масштабу (Scale Factor) (а), мультискейлінгова поверхня нечіткої ентропії (б) та усереднена по всім масштабам нечітка ентропія (в).

На рис. 3.5 видно, як поводить себе мультимасштабна нечітка ентропія під час крахових явищ: вона спадає у перед кризовий стан та зростає під час криз. Гранульована поверхня показує, що навіть на більших масштабах спостерігаються стрибки даного показника під час криз, що може слугувати в якості індикатора кризової події.

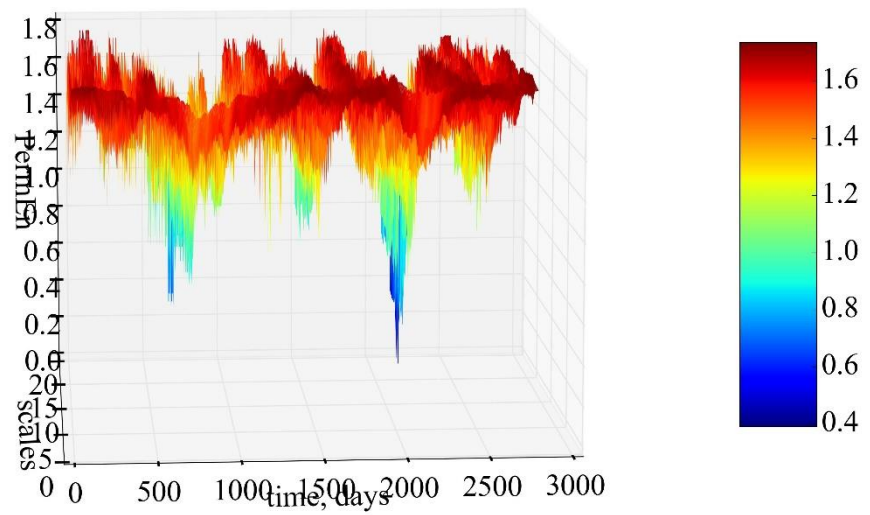
### 3.4 Пермутаційна ентропія

На рис. 3.6 представлено мультимасштабну пермутаційну ентропію для всього часового ряду, з врахуванням перших 20 масштабів та на основі алгоритму рухомого вікна.

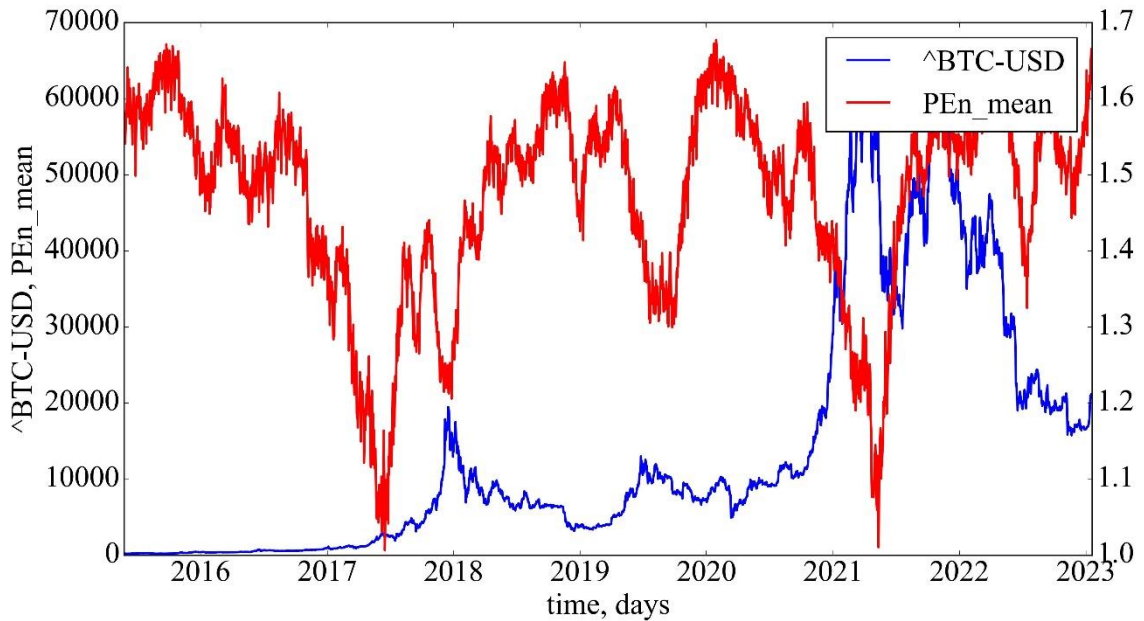




(a)



(b)



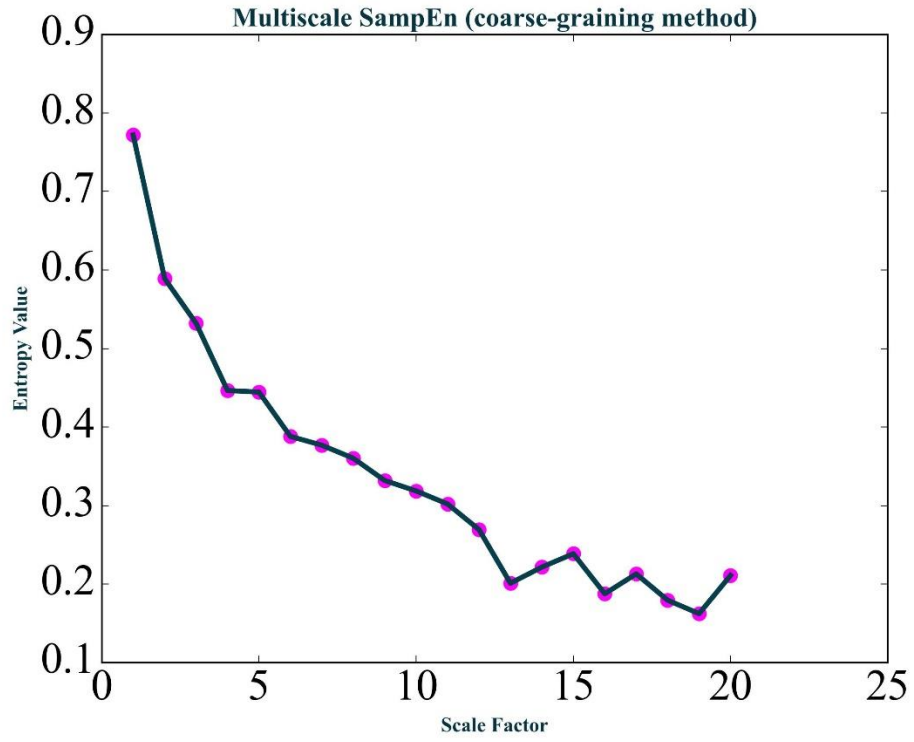
(в)

**Рис. 3.6.** Залежність пермутаційної ентропії (Entropy Value) від масштабу (Scale Factor) (а), мультискейлінгова поверхня пермутаційної ентропії (б) та усереднена по всім масштабам пермутаційна ентропія (в).

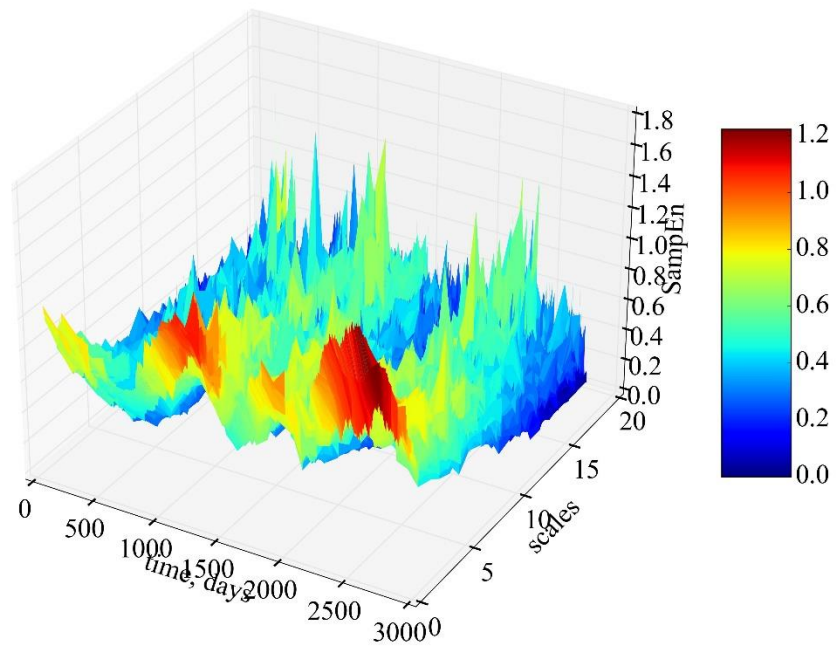
На рис. 3.6 видно, як спадає пермутаційна ентропія у передкризовий стан. Подібна динаміка вказує на зростання кількості однотипних шаблонів, що відповідає спаду даного показника на всіх часових масштабах.

### 3.5 Ентропія шаблонів

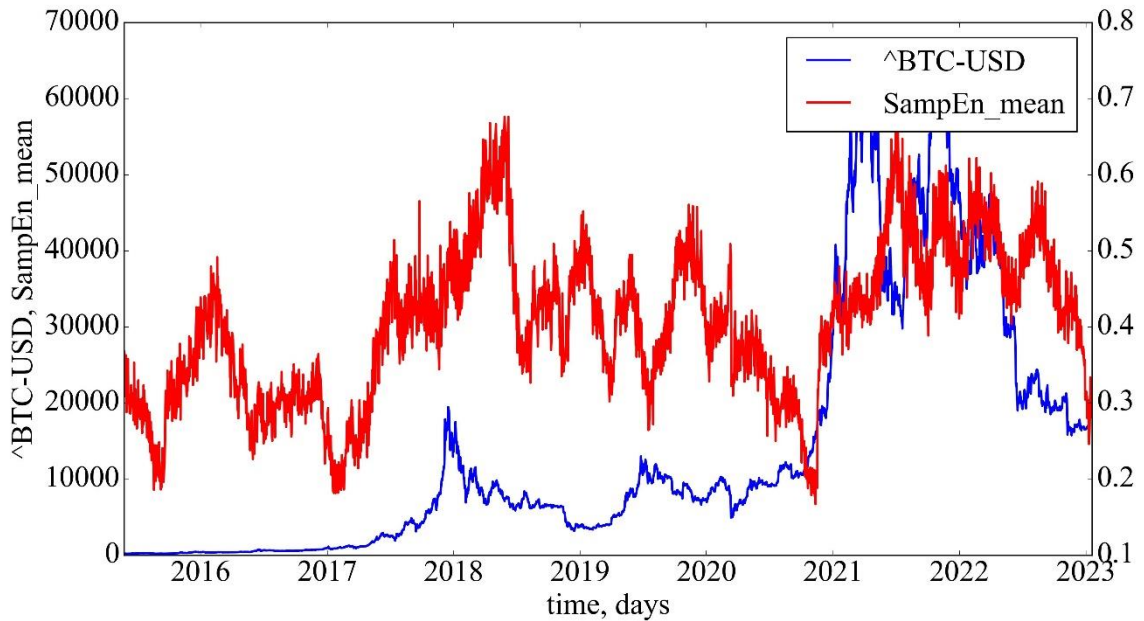
На рис. 3.7 представлено мультимасштабну ентропію шаблонів для всього часового ряду, з врахуванням перших 20 масштабів та на основі алгоритму рухомого вікна.



(a)



(b)



(в)

**Рис. 3.7.** Залежність ентропії шаблонів (Entropy Value) від масштабу (Scale Factor) (а), мультискейлінгова поверхня ентропії шаблонів (б) та усереднена по всім масштабам ентропія шаблонів (в).

З рис. 3.7 видно, що даний показник характеризується ростом на всіх часових проміжках під час кризового стану, вказуючи на особливу чутливість ринку в даний період. Також ця закономірність чітко просліджується, якщо подивитися на ентропію шаблонів усереднену по всім часовим масштабам.

## ВИСНОВКИ

Таким чином, певні аспекти складності економічної системи можна вивчати, розглядаючи інформацію про ступінь складності. Особливо ефективними є різномасштабні варіації впроваджуваної Шенонівської ентропії.

Ретельно аналізуючи систем різної природи й складності з використанням показників теорії складності, стає можливим зрозуміти їх природу і передбачити потенційно критичну поведінку системи.

Співставляючи часовий ряд біткоїна з відповідними індексами складності, реакція на критичні явища очевидна. Однак використання їх як передвісників криптовалютних криз є проблемою, яка потребує подальшого дослідження. Це дослідження підкреслює переваги використання показників мультимасштабних ентропій. Метод грануляції ряду є дуже універсальним і демонструє високий ступінь чутливості до крахів при різних часових масштабах.

У подальшому цей метод можна адаптувати до систем підтримки прийняття рішень для захисту приватних інвестицій та зниження потенційних інвестиційних ризиків.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Investopedia.com What is bitcoin?  
<https://www.investopedia.com/terms/b/bitcoin.asp>
2. B.M. Cole, A.H. Dyhrberg, S. Foley, J. Svec, Can Bitcoin be Trusted? Quantifying the Economic Value of Blockchain Transactions, Journal of International Financial Markets, Institutions & Money (2022), doi:  
<https://doi.org/10.1016/j.intfin.2022.101577>
3. John E. Marthinsen, Steven R. Gordon. The price and cost of bitcoin, The Quarterly Review of Economics and Finance, Volume 85, 2022, Pages 280-288, ISSN 1062-9769, <https://doi.org/10.1016/j.qref.2022.04.003>
4. Khaled Mokni, Ahmed Bouteska, Mohamed Sahbi Nakhli. Investor sentiment and Bitcoin relationship: A quantile-based analysis, North American Journal of Economics and Finance, 60, 2022, 101657
5. Vault12.com What is entropy? Entropy helps you're your digital assets safe <https://www.investopedia.com/terms/b/bitcoin.asp>
6. Delgado-Bonal, A., & Marshak, A. (2019). Approximate Entropy and Sample Entropy: A Comprehensive Tutorial. Entropy (Basel, Switzerland), 21(6), 541, <https://doi.org/10.3390/e21060541>
7. Bandt, C., & Pompe, B. (2002). Permutation entropy: a natural complexity measure for time series. *Physical review letters*, 88(17), 174102, <https://doi.org/10.1103/PhysRevLett.88.174102>
8. M. Rostaghi and H. Azami, Dispersion Entropy: A Measure for Time-Series Analysis, in IEEE Signal Processing Letters, vol. 23, no. 5, pp. 610-614, May 2016, <https://doi.org/10.1109/LSP.2016.2542881>
9. Hong-Bo Xie, Wei-Xing He, Hui Liu, Measuring time series regularity using nonlinear similarity-based sample entropy, Physics Letters A, Volume 372, Issue 48, 2008, Pages 7140-7146, ISSN 0375-9601, <https://doi.org/10.1016/j.physleta.2008.10.049>