УДК 004.75:519.8 DOI: 10.30837/0135-1710.2025.186.071

Є.Є. ДЕМЕНКО, І.В. ГРЕБЕННІК, М.М. КОЛМИКОВ

МЕТОД СТВОРЕННЯ ДАТАСЕТІВ ДЛЯ ОЦІНКИ АЛГОРИТМІВ РОЗПОДІЛУ ВАЛІДАТОРІВ НА ОСНОВІ МЕХАНІЗМУ PROOF OF STAKE

Досліджено проблему відтворюваності експериментів при оптимізації розподілу валідаторів у блокчейн-мережах із консенсусом Proof of Stake (PoS), зокрема через відсутність стандартизованих датасетів та уніфікованих методів тестування, що ускладнює об'єктивне порівняння алгоритмів. Для вирішення цієї проблеми запропоновано метод побудови тестових наборів даних із детермінованими генераторами псевдовипадкових послідовностей та характеристиками валідаторів, налаштованими за статистикою мережі Ethereum 2.0, включно з розміром стейку, продуктивністю, надійністю, мережевими затримками та географією. Створено три набори даних різного масштабу (50, 500, 1000 валідаторів) із фіксованими seed-значеннями для повної відтворюваності експериментів, розроблено систему критеріїв оцінки та порядок виконання тестів для підвищення статистичної достовірності. У експериментах оцінено чотири алгоритми розподілу: PSO+LS, Random+Repair, EthShuffle та Greedy. Результати показали масштабно-залежну ефективність: PSO+LS забезпечує високу якість оптимізації, але значні обчислювальні витрати обмежують його застосування офлайн-плануванням; EthShuffle має найменшу варіативність при середній ефективності; Random+Repair швидший, але менш стабільний; Greedy демонструє максимальну швидкість із детерміністичними результатами, проте змінною ефективністю при масштабуванні. Запропоновані набори даних та методи створюють основу для об'єктивного порівняння нових алгоритмів PoS-систем та підвищення надійності управління мережею.

1. Вступ

Еволюція технологій розподіленого реєстру відзначається переходом від енергоспоживних консенсусних механізмів до ефективніших. Серед провідних блокчейн-платформ першість належить саме механізму Proof of Stake (PoS) завдяки значному скороченню споживання енергії, безпека та децентралізація зберігаються на високих рівнях. Класичний приклад успішного впровадження PoS можна спостерігати в мережі Ethereum 2.0, яка була запущена в 2022 році і продемонструвала практичну життєздатність цього підходу [1].

Ключовою особливістю PoS-систем є необхідність ефективного розподілу валідаторів між комітетами та шардами для забезпечення оптимального функціонування мережі. У контексті Ethereum 2.0 цей процес включає координацію роботи понад тисяч валідаторів через Beacon Chain, формування комітетів розміром 128 учасників та їх розподіл між 64 потенційними шардами [2], [3]. Алгоритм RANDAO забезпечує псевдовипадковий вибір валідаторів кожні 12 секунд для створення та підтвердження блоків, при цьому механізм shuffling гарантує регулярну ротацію учасників між різними ролями [4].

Ефективність алгоритмів розподілу валідаторів забезпечує реалізацію трьох ключових засад блокчейн-систем: масштабованості, децентралізації та безпеки. Хоча науковці активно займаються розробками нових алгоритмів, однак питання коректного тестування та порівняння їх залишається недостатньо вивченим. Це призводить до складності: різні дослідники використовують різні методи тестування та різні набори даних і об'єктивно визначити найкращий алгоритм стає практично неможливо [5].

Ситуація ускладнюється тим, що чимало вчених не оприлюднюють здобуті експериментальні дані, не надають розгорнутого опису використаних методик збирання цих даних. Така непрозорість помітно гальмує прогрес у дослідницькій галузі [5]. Проблема ускладнює порівняння різних алгоритмів і перевірку їхніх результатів.

Окремим викликом виступає масштабованість досліджень. Алгоритми, що демонструють відмінні результати на невеликих тестових наборах даних, можуть показувати кардинально іншу поведінку при збільшенні розміру мережі до реальних параметрів. Мережа Ethereum 2.0 продовжує стрімко розвиватися [6]. Таке швидке масштабування висуває додаткові вимоги до алгоритмів розподілу, які мають ефективно функціонувати в умовах постійного зростання кількості учасників.

Практичне значення проблеми підкреслюється зростаючою роллю PoS-систем у критичній інфраструктурі. Децентралізовані фінансові сервіси, побудовані на основі PoS-блокчейнів, обробляють транзакції на мільярди доларів щодня [7]. Державні установи розглядають можливості впровадження блокчейн-технологій для реалізації електронного урядування та цифрових валют центральних банків. Таким чином, у цьому контексті надійність і продуктивність алгоритмів розподілу валідаторів мають подвійне значення. Їхнє значення виходить за межі простої оптимізації. Правильний розподіл валідаторів важливий і для запобігання масовим відмовам, особливо у високонавантажених або критичних системах, таких як блокчейн-платформи та децентралізовані фінансові мережі.

Цінність цього дослідження полягає в прагненні до максимальної відтворюваності. Формування стартових наборів даних дає кілька переваг. По-перше, підвищується точність результатів. По-друге, пришвидшується розвиток галузі. По-третє, створюється міцна основа для впровадження рішень у реальні системи.

2. Аналіз сучасних наукових публікацій і постановка проблеми дослідження

Сфера розподілу валідаторів у PoS є великою й різноплановою. Вона охоплює велику кількість алгоритмів, що поєднують криптографічні засади з реальними потребами розподілу. Існуючі дослідження ґрунтовно висвітлюють теоретичні основи цих алгоритмів [1], [4], [5]. Проте в процесі їх практичного впровадження підходи до експериментальної оцінки алгоритмів у різних дослідженнях суттєво різняться [8].

Процес випадкового вибору валідаторів необхідний для PoS-механізмів консенсусу, оскільки непередбачуваність може тільки забезпечити достатню безпеку, аби зберегти пропорційне відображення на основі ваги стейків [9]. Однак PoS-механізми Ethereum стають повноцінними, незважаючи на це, завдяки алгоритму RANDAO, оскільки випадковість, необхідна для обчислення та перетасування валідаторів, є ключовою функціональністю [4]. Механізм RANDAO в Ethereum реалізується з використанням хеш-функції Keccak-256, seed-значень із попередніх блоків і верифікованих випадкових функцій, що забезпечують прозорість та непередбачуваність системи.

Технічно складнішим є алгоритм Swap-or-Not Shuffle, адаптований із криптографічних досліджень перетасування карт [10]. Алгоритм реалізується через серію умовних обмінів у 90 раундах, що базуються на безпековому аналізі приблизно $4\log_2(N)$ раундів для N валідаторів [11]. Основна перевага полягає в його властивості незалежності від вхідних значень, що дозволяє обчислювати перетасування для окремих індексів без необхідності перемішувати весь список валідаторів, що критично для підтримки легких клієнтів. Алгоритм забезпечує O(1) обчислень для визначення пункту призначення окремого індексу та O(1) зворотне обчислення без потреби зберігати або обробляти повний набір валідаторів. Така характеристика робить Swap-or-Not Shuffle особливо ефективним для масштабних блокчейн-систем, де повне перетасування є обчислювально затратним.

Перехід від класичних алгоритмів до метаевристик відкриває нові можливості для оптимізації розподілу валідаторів у PoS-системах. Водночас відносно мало досліджень присвячено застосуванню таких методів саме для цієї задачі. Застосування генетичних алгоритмів у сфері блокчейну стосуються гібридного консенсусу та покращень блокчейну в цілому, а не конкретно проблеми вибору валідаторів [12]. Це свідчить про область, яка

ще недостатньо досліджена. У ній метаевристичні методи здатні продемонструвати відчутний прогрес.

Particle Swarm Optimization (PSO) показує обнадійливіші результати у контексті блокчейн оптимізації. Дослідження демонструють Multi-Objective Particle Swarm Optimization (MOPSO) застосування для оптимізації конфігурації блокчейну, показуючи що Парето-оптимальні конфігурації можуть бути знайдені за допомогою PSO підходів [13]. Багатоцільова оптимізація розглядає вибір валідаторів як багатоцільову проблему, враховуючи безпеку, продуктивність, децентралізацію та енергоефективність. Ці дослідження показують перспективу створення гібридних алгоритмів. Вони можуть поєднати сильні сторони метаевристик із класичними методами розподілу валідаторів.

Існуючі методології експериментальної оцінки мають серйозні обмеження. Вони стають особливо помітними в сучасних дослідженнях PoS-систем. У наукових роботах застосовуються різні підходи – від симуляційних експериментів до аналітичних фреймворків. Зокрема, PoS-симулятор, реалізований на мові Julia [8], та AlphaBlock Framework [14] пропонують спеціалізовані інструменти для моделювання та тестування різних варіантів PoS. Втім, відсутність уніфікованих стандартів щодо їх застосування унеможливлює об'єктивне порівняння отриманих результатів та знижує відтворюваність досліджень.

Метрики оцінки продуктивності суттєво варіюють між різними дослідженнями, що ускладнює об'єктивне порівняння алгоритмів. До первинних метрик належать пропускна здатність, затримка, час фіналізації та споживання енергії. Економічні метрики охоплюють коефіцієнт Джині для оцінки рівності розподілу багатства, структуру розподілу винагород та рівень централізації валідаторів [5]. Безпекові метрики, в свою чергу, зосереджені на стійкості до атак, зокрема економічній вартості атаки 51 %, ефективності механізмів штрафів для проблеми nothing-at-stake та стійкості до long-range атак. Відсутність консенсусу щодо відносної важливості цих метрик призводить до фрагментованості дослідницького ландшафту.

Фундаментальна проблема відтворюваності окреслена дослідниками Algorand і полягає у відсутності уніфікованого фреймворку для аналізу продуктивності блокчейн-систем [15]. На практиці різні проєкти застосовують несумісні метрики та експериментальні умови, що робить змістовне порівняння результатів складним. У публікаціях нерідко наводяться показники у сотні, тисячі чи навіть мільйони транзакцій на секунду, однак базові припущення й параметри експериментів залишаються неявними. Дослідники Algorand виокремили сім критичних вимірів, які суттєво впливають на оцінку продуктивності блокчейну і водночас варіюються неконсистентно між роботами: режим доступу (permissioned або permissionless), модель участі (пряма або делегована), тип мережевої інфраструктури, рівень видимості транзакцій, профіль розподілу пропускної здатності, прийнята модель безпеки та використані механізми стиснення.

Дослідження BFT-консенсусу нерідко дають різні результати. Іноді вони навіть суперечать одне одному, особливо у питаннях масштабованості [16]. Причина полягає у відмінностях умов проведення експериментів. Це робить їхнє пряме порівняння складним.

У випадку з алгоритмами розподілу валідаторів у Proof of Stake ситуація ще складніша. Дослідження у цій сфері залишаються фрагментованими. У методології існують значні прогалини, які непросто усунути. Порівняти результати між роботами майже неможливо, адже кожне дослідження застосовує власні метрики, відмінні умови тестування та специфічні реалізації. Моделі й мережі теж можуть відрізнятися значно. Лише комплексний підхід здатен змінити ситуацію. Залишається нагальна потреба в уніфікації метрик, відтворюваних процедур та підвищенні наукової цінності отриманих результатів.

3. Мета і задачі дослідження

Головною метою цього наукового дослідження є розробка методу створення експериментальних датасетів, за допомогою яких можна здійснювати оцінку ефективності алгоритмів розподілу валідаторів у блокчейн-системах на конкретному прикладі роботи механізму консенсусу Proof of Stake. Успішна реалізація дозволить усунути значну проблему відсутності відтворюваних підходів для експериментальної перевірки існуючих алгоритмів розподілу валідаторів і знайти можливість об'єктивно порівняти їх. Цей результат можливо досягти за рахунок вирішення таких задач:

– розробка формального методу створення реалістичних параметрів валідаторів, що враховує важливі аспекти PoS-систем, зокрема розподіл стейків, надійність учасників і мережеві затримки;

– побудова трьох типів експериментальних датасетів різного масштабу (50, 500, 1000 валідаторів) з фіксованими параметрами генерації для забезпечення відтворюваності результатів;

– визначення комплексних критеріїв оцінки алгоритмів розподілу валідаторів, які включають метрики якості розподілу, стабільності результатів та обчислювальної ефективності;

– проведення системних експериментальних досліджень для оцінки ефективності базових та гібридних алгоритмів на розроблених датасетах і отримання порівняльних характеристик продуктивності;

– аналіз масштабованості алгоритмів при зростанні розміру мережі валідаторів для подальшого формулювання науково обґрунтованих рекомендацій щодо їх практичного застосування у реальних PoS-системах.

4. Метод побудови експериментальних датасетів

Через потребу у подоланні фрагментарності підходів до реалізації експериментальної оцінки алгоритмів перерозподілу валідаторів виникла необхідність у формально обґрунтованому методі генерування даних, на основі якої формуються експериментальні набори.

Кожен i-й валідатор у системі характеризується набором параметрів:

$$V_i = (s_i, p_i, r_i, l_{ij}, g_i, q_i, h_i), \tag{1}$$

де s_i – розмір стейку; p_i – продуктивність; r_i – надійність; l_{ij} – мережеву затримку між i-им та j-им валідаторами; g_i – географічне розташування; q_i – якість мережевого з'єднання; h_i – історія штрафів та порушень протоколу.

Для забезпечення реалістичності моделювання кожен валідатор має мінімальний стейк 32 ETH, що відповідає правилам Ethereum 2.0 [2]. Такий підхід дозволяє точно відобразити механіку роботи валідаторів і уникнути спотворення результатів через нереалістично великі значення стейку. Інші параметри валідаторів, такі як продуктивність, надійність та мережеві затримки, генеруються у реалістичних діапазонах, що відповідають спектру реальних учасників мережі Ethereum 2.0 [3]. Продуктивність валідаторів генерується як:

$$p_i \sim U(0.5,1.5) \cdot p_{\text{base}}\,, \tag{2}$$

де p_{base} – базовий рівень продуктивності. Розподіл відображає різноманітність технічних характеристик обладнання валідаторів. Діапазон від 0.5 до 1.5 обрано на основі емпіричного аналізу статистики Beacon Chain у мережі Ethereum 2.0 [3]. Поточна продуктивність валідаторів оцінюється через ефективність участі у консенсусі, що включає три складові: ефективність атестацій (attester efficiency), ефективність пропозицій блоків (proposer

efficiency) та ефективність участі у синхронізаційних комітетах (sync committee efficiency), які відповідно складають 84,4 %, 12,5 % та 3,1 % від винагород валідатора [17]. Загальна ефективність обчислюється як відношення фактичної винагороди до ідеальної, що дозволяє врахувати відхилення продуктивності від середнього рівня. Нижня межа 0.5 відповідає мінімально допустимій продуктивності для участі у консенсусі, тоді як верхня межа 1.5 характеризує високопродуктивні вузли з оптимізованими конфігураціями.

Надійність валідаторів моделюється у вигляді:

$$r_i \sim U(0.8,1.0). \tag{3}$$

Межі надійності встановлені згідно з аналізом історії штрафів у Beacon Chain [2, 5]. Валідатори з надійністю нижче 80% мають підвищений ризик slashing через пропуски атестацій. Це значення відповідає практичному порогу надійності для стабільної роботи в PoS-мережах.

Мережеві затримки формуються як симетрична матриця $L = [l_{ij}]_{n\times n}$:

$$l_{ij} = \begin{cases} U(10,50)\text{ мс}, & g_i = g_j \text{ (той самий регіон)} \\ U(50,200)\text{ мс}, & g_i \neq g_j \text{ (різні регіони)} \end{cases}, \tag{4}$$

де g_i та g_j – географічні регіони *i*-го та *j*-го валідаторів відповідно. Така структура забезпечує реалістичне моделювання глобальної топології мережі з урахуванням географічного розподілу учасників. Діапазони мережевих затримок базуються на вимірах реальних P2P-мереж блокчейнів. Внутрішньорегіональні затримки 10-50 мс та міжрегіональні 50-200 мс відповідають типовим показникам для забезпечення 12-секундного циклу створення блоків у Ethereum 2.0 [18]. Географічне розташування валідаторів моделюється шляхом їх розподілу між основними регіонами на основі актуальних досліджень мережі Ethereum. Згідно з [19], фактичний розподіл має такий вигляд: Європа – 46.79 %, Північна Америка – 38.05 %, Азія – 9.95 %, Океанія – 4.16 %.

Якість мережевого з'єднання валідаторів генерується в діапазоні від 10 до 1000 Мбіт/с для моделювання різноманітності учасників мережі. Мінімальна вимога 10 Мбіт/с забезпечує передачу даних у межах 12-секундного циклу Ethereum 2.0 [6]. Розподіл налаштовано так, щоб більшість валідаторів мала середні показники, а меншість – високі, що відповідає типовій структурі децентралізованих мереж з різнорідними учасниками.

Історія штрафів валідаторів моделюється як накопичувальний показник порушень протоколу з урахуванням часового фактору. Система враховує чотири типи подій: відсутність порушень, пропуски атестацій, некоректні атестації та найсерйозніше порушення – подвійне підписання блоків. Недавні події мають більшу вагу при розрахунку загального показника, що відповідає принципам репутаційних систем у блокчейні. Частота порушень визначається згідно зі статистикою Beacon Chain [2], де приблизно 2 % валідаторів мають зафіксовані порушення протягом 100 епох. Валідатори з високим рівнем накопичених штрафів класифікуються як потенційно ненадійні для цілей Byzantine Fault Tolerance [16], що впливає на їх розподіл між комітетами для забезпечення безпеки мережі.

Для гарантування повної відтворюваності експериментів встановлено строгий протокол фіксації початкових значень генераторів псевдовипадкових чисел. Кожному набору даних призначається унікальний ідентифікатор $seed(n, version)$, що включає розмір набору та версію генератора. Всі згенеровані датасети зберігаються у форматі з метаданими про параметри генерації та контрольними сумами для верифікації цілісності.

Було сформовано три групи наборів даних, які покривають широку палітру практичних прикладів підходу до системи. Сценарій з малою кількістю валідаторів

(приблизно 50) моделював приватні та консорціумні блокчейн-системи. Середній із 500 валідаторів відображав підходи до корпоративних систем блоку або тестових мереж та забезпечував співвідношення рівності обчислення, якісності і доступності. Набір даних із 1000 валідаторів використовувався для моделювання роботи алгоритмів у міру збільшення масштабу та відображав характерні особливості публічних блокчейн-мереж. Хоча ця кількість значно менша за реальну мережу Ethereum, вона все ж дозволила вивчати поведінку алгоритмів у порівняно великих підмережах. Це дало змогу оцінити їхню масштабованість.

Крім того, слід обирати значення seed так, щоб забезпечити статистичну незалежність, псевдовипадковість, передбачуваність і масштабованість створюваних серій наборів даних. Ці значення не можуть бути кратними або степенями двійки. Рекомендується псевдовипадковий інтервал, який не є надто близьким до будь-якого з іншого цілого числа, і перевіряти незалежність за допомогою коефіцієнтів кореляції. Seed можна обирати послідовно, за допомогою хешування параметрів набору даних або у ієрархічній формі, що враховує категорії експерименту, розмір набору даних та індекс екземпляра.

Перевірка seed передбачає спочатку оцінку рівномірності розподілу даних, а потім аналіз кореляцій та повторюваних патернів. Такий підхід закладає міцну основу для подальшого розширення датасетів і проведення нових експериментів. Він має особливе значення для забезпечення відтворюваності результатів у PoS-системах.

Для всебічної оцінки алгоритмів визначено систему комплексних критеріїв, що охоплює якість розподілу, стабільність результатів та обчислювальну ефективність. Критерієм якості є значення узагальненого критерію Z [20]:

$$Z = \alpha \cdot F_1(X) - \beta \cdot F_2(X) - \gamma \cdot F_3(X)\,, \tag{5}$$

де $F_1(X)$ максимізує пропускну здатність системи, $F_2(X)$мінімізує дисбаланс навантаження, $F_2(X)$ мінімізує мережеві затримки. Значення Z дозволяє визначати оптимальних валідаторів для включення до комітету: валідатор зараховується, якщо його Z перевищує встановлений поріг або є найвищим серед кандидатів. Значення вагових коефіцієнтів $\alpha = 1.0$, $\beta = 0.3$, $\gamma = 0.2$ встановлено на основі аналізу подібних багатокритеріальних задач оптимізації в розподілених системах. Співвідношення узгоджується з рекомендаціями для балансування продуктивності та надійності в консенсусних протоколах, де пропускна здатність має найвищий пріоритет [20]. Додатково проведено валідацію через серію експериментів з альтернативними конфігураціями коефіцієнтів, які підтвердили оптимальність обраних значень для досліджуваних масштабів мережі.

Стабільність алгоритмів оцінюється через стандартне відхилення $\sigma(Z)$, обчислене за серією з $N = 10$ незалежних запусків:

$$\sigma(Z) = \sqrt{\frac{1}{N-1}\sum_{k=1}^{N}(Z_k - \bar{Z})^2}\,, \tag{6}$$

де Z_k – результат k-го запуску.

Обчислювальна ефективність оцінюється за двома показниками: часом виконання $T(n)$ та кількістю обчислень узагальненого критерію. Для дослідження масштабованості використовується метрика:

$$S(n) = \frac{T(n)}{T(n_0)}\,, \tag{7}$$

де $T(n)$ – час виконання задачі розміру n; $T(n_0)$ – базовий час виконання для референтного розміру задачі n_0. Ця метрика відображає, наскільки зростає час виконання із збільшенням розміру задачі, і дозволяє кількісно оцінити масштабованість алгоритмів.

Розроблено протокол для системного оцінювання ефективності алгоритмів розподілу валідаторів, який передбачає тестування чотирьох алгоритмів:

– Гібридний PSO+LS — поєднує метод рою частинок з локальним пошуком [20], [21];

– Random+Repair — базовий алгоритм із випадковою генерацією та корекцією;

– Greedy — жадібний алгоритм [22];

– EthShuffle — адаптація механізму RANDAO з Ethereum 2.0 [4].

Кожен алгоритм було протестовано на всіх наборах даних із усіма метриками десять разів. Такий підхід забезпечив статистичну надійність результатів, зменшуючи вплив випадковості та дозволяючи оцінити дисперсію, при цьому не створюючи надмірних обчислювальних витрат. Попереднє тестування показало, що стандартне відхилення результатів стабілізується після 8-10 запусків, і додаткові запуски не впливають на висновки щодо відносної ефективності алгоритмів. Отже, десятикратне повторення експериментів є оптимальним компромісом між достовірністю результатів і обчислювальною ефективністю. Тому кожен алгоритм було протестовано на кожному наборі даних із усіма метриками десять разів, що становить оптимальний компроміс між достовірністю й ефективністю.

Тести було проведено з використанням Python v3.9. Для роботи з матрицями було використано NumPy, а паралельні обчислення організовано через multiprocessing. Результати було збережено у вигляді ієрархічної таблиці, що допомагає легше аналізувати дані.

5. Результати дослідження

Перед проведенням основних експериментів було проаналізовано характеристики згенерованих наборів даних для підтвердження їх реалістичності та відтворюваності. Гістограми розподілу стейків (рис. 1) свідчать про логнормальний характер розподілу. Для малого набору (n=50) спостерігається більша варіативність розподілу, тоді як для великих наборів (n=500, n=1000) розподіл стає стабільнішим.

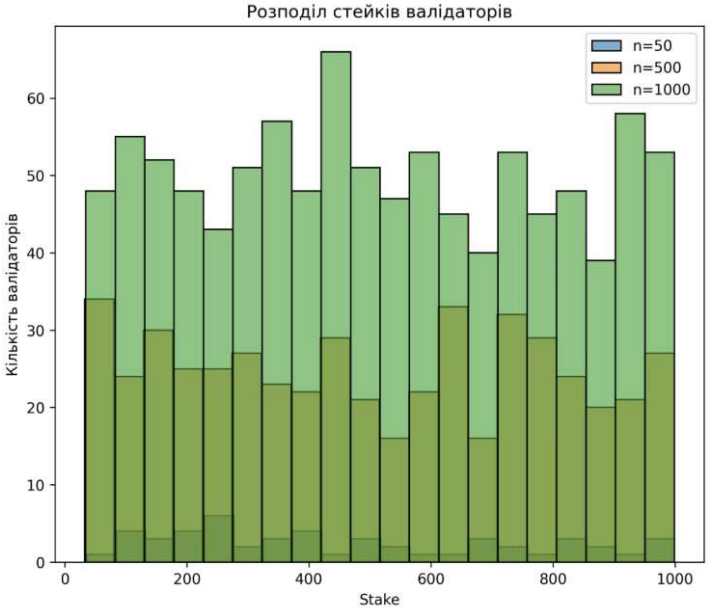

Рис. 1. Розподіл стейків валідаторів для трьох масштабів

На діаграмі розсіювання, наведеній на рис. 2, відображено залежності якості рішень від часу виконання чотирьох алгоритмів розподілу валідаторів у PoS-системах. PSO+LS забезпечує високу якість оптимізації серед досліджуваних алгоритмів, однак за рахунок значно більшого часу виконання. Random+Repair показує помірну швидкість виконання з варіативними результатами. EthShuffle демонструє середню ефективність при швидкому виконанні. Алгоритм Greedy вирізняється високою швидкістю виконання та стабільними детерміністичними результатами, однак ефективність його рішень суттєво залежить від масштабу задачі. Використання логарифмічної шкали часу демонструє баланс між ефективністю та швидкістю виконання. Такий підхід дозволяє наочно оцінити компромісні рішення.

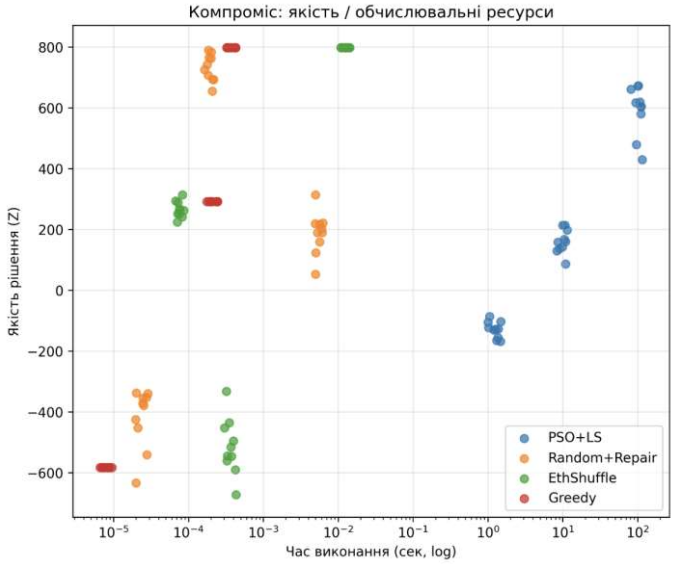

Рис. 2. Діаграма компромісу «якість/обчислювальні ресурси»

Квадратичне зростання часу виконання PSO+LS робить його придатним для офлайн-планування розподілу валідаторів, де якість оптимізації важливіша за швидкість. Базові алгоритми демонструють лінійну або сублінійну складність, що робить їх придатнішими для задач реального часу, але з суттєвою втратою якості оптимізації.

Матриця кореляції між результатами виконання алгоритмів (рис. 3) демонструє використання фіксованих seed-значень для забезпечення відтворюваності експериментів. Алгоритм Greedy не має кореляційних значень (білі комірки), оскільки він є детерміністичним і завжди повертає однаковий результат. Для інших пар алгоритмів коефіцієнти кореляції перебувають у діапазоні від –0.11 до 0.03, а відповідні p-value перевищують 0.05, що вказує на відсутність статистично значущих лінійних зв'язків між результатами різних алгоритмів. Низькі коефіцієнти кореляції між результатами алгоритмів відповідають очікуванням для методів з різними принципами оптимізації.

Аналіз кумулятивної функції розподілу (CDF) якості рішень (див. рис. 4) показує помітні відмінності у поведінці алгоритмів. Зокрема, алгоритм PSO+LS характеризується найширшим діапазоном отриманих результатів, що свідчить про його здатність охоплювати різні траєкторії пошуку та знаходити рішення різної якості. Метод Random+Repair займає проміжне положення: він має вищу стабільність порівняно з PSO+LS, однак поступається йому за різноманітністю результатів. Алгоритм EthShuffle демонструє стрімкіший ріст кумулятивної кривої, що вказує на меншу варіативність

результатів і концентрацію рішень у відносно вузькому інтервалі якості. Найобмеженішою є поведінка алгоритму Greedy: крива поведінки має характерні стрибки та демонструє мінімальну варіативність, що узгоджується з його детерміністичною природою.

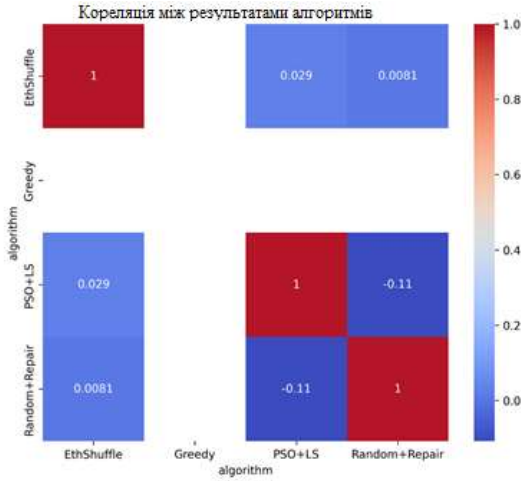

Рис. 3. Матриця кореляції між результатами алгоритмів

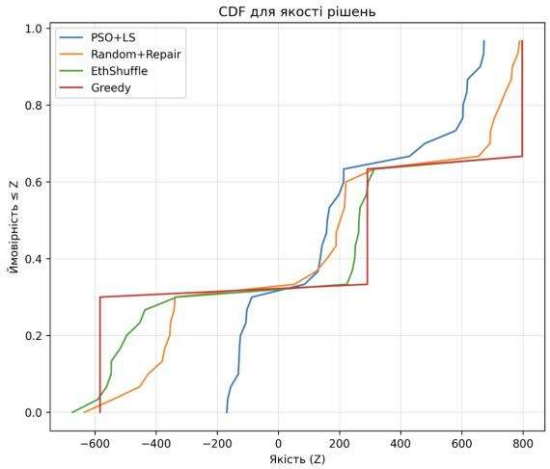

Рис. 4. Кумулятивна функція розподілу (CDF) якості рішень

6. Обговорення результатів дослідження

За результатами дослідження представлено метод формування експериментальних даних, який забезпечує відтворюваність при дослідженні алгоритмів розподілу валідаторів у PoS-системах. Раніше запропоновані підходи часто не містили достатньо деталей про умови проведення експериментів. На їх основі створювалися нові алгоритми, проте результати важко було порівнювати. Використання запропонованого методу дозволяє виконувати серію експериментів із заздалегідь визначеними параметрами, що гарантує стабільність і порівнянність результатів.

Оцінювання алгоритмів здійснювалося за кількома показниками: ефективність розподілу, стабільність роботи та обчислювальні витрати. Такий підхід дозволив отримати повне уявлення про їхні сильні й слабкі сторони та зробити обґрунтований вибір для конкретних сценаріїв застосування.

Результати експерименту виявили особливу поведінку алгоритмів на різних масштабах мережі. Ефективність алгоритму Greedy залежить від масштабу: на середніх масштабах мережі він показує конкурентоспроможні результати, тоді як на малих наборах даних його продуктивність знижується. Така поведінка пояснюється детерміністичною природою алгоритму, який завжди приймає локально оптимальні рішення без урахування глобального контексту. Завдяки масштабованому аналізу на трьох рівнях (50, 500, 1000 валідаторів) стали видимими нелінійні залежності продуктивності алгоритму від розміру мережі. Наприклад, PSO+LS демонструє високу якість оптимізації на всіх масштабах мережі, проте квадратичне зростання часу виконання обмежує його застосування виключно сферою офлайн-планування, де час не є критичним фактором.

Статистична валідація згенерованих датасетів підтвердила основні характеристики синтетичних даних. Коефіцієнт Джині [5] для розподілу стейків складає 0.31±0.01, що свідчить про помірну нерівність, характерну для децентралізованих мереж. Мережеві параметри показали реалістичні значення: латентність між валідаторами одного регіону становить 30±11 мс, між різними регіонами – 125±43 мс, що відповідає типовим показникам глобальних P2P-мереж. Частка потенційно ненадійних валідаторів (24-25 %)

міститься в безпечному діапазоні для Byzantine Fault Tolerance (< 33 %). Географічний розподіл валідаторів рівномірно покриває чотири регіони з варіацією 18-24 %, що забезпечує достатню децентралізацію. Слабка кореляція між стейком та продуктивністю ($|r| < 0.15$) відображає реалістичну незалежність економічних можливостей учасників від їхніх технічних характеристик, що є типовим для відкритих PoS-систем, де розмір стейку не гарантує якість обладнання.

Разом із перевагами важливо окреслити і межі застосовності. Експериментальні дані отримано не з реальної мережі, а згенеровано програмно. Для побудови використано статистичні параметри, виділені з аналізу існуючих PoS-систем. В основу покладено емпіричні характеристики Ethereum 2.0, що дозволило сформувати синтетичні датасети з високим рівнем реалістичності. Водночас збережено повну відтворюваність експериментів, що забезпечує можливість їх перевірки та повторення. Синтетичні датасети відтворюють статичні характеристики валідаторів (розподіл стейків, продуктивність, надійність, географічне розташування), однак мають обмеження у відтворенні динамічних аспектів реальних мереж. Зокрема, не враховуються часові залежності поведінки валідаторів, економічні стимули, що змінюються, та адаптивні стратегії учасників мережі. Крім того, у дослідженні використано 1000 валідаторів, що є спрощенням у порівнянні з масштабами мереж на кшталт Ethereum 2.0, однак цього обсягу достатньо для перевірки коректності та працездатності запропонованого методу.

Що стосується обчислень, PSO+LS вимагає значних ресурсів і часу виконання, які зростають зі збільшенням розміру мережі. Для офлайн-експериментів це прийнятно, але для систем із швидкою реконфігурацією комітетів може бути проблемою. Таке обмеження робить PSO+LS непридатним для адаптації в режимі реального часу. Алгоритм доцільно використовувати для довгострокового планування конфігурації комітетів або періодичної оптимізації розподілу між епохами.

Попри зазначені обмеження синтетичних даних, експериментальні результати дозволяють сформулювати практичні рекомендації для різних сценаріїв застосування алгоритмів у реальних PoS-системах. Виявлені закономірності поведінки алгоритмів на різних масштабах мережі створюють основу для їх цільового використання відповідно до специфічних вимог системи. На основі результатів сформульовано практичні рекомендації: для критичних застосувань, де можливе періодичне офлайн-планування конфігурації (наприклад, раз на добу або між великими оновленнями мережі), рекомендується PSO+LS; для систем реального часу, що потребують балансу між якістю та швидкістю – EthShuffle; для швидкої реконфігурації мереж – Random+Repair, для простих систем з обмеженими ресурсами, де швидкість критична, а оптимальність вторинна, – Greedy.

У майбутніх дослідженнях варто зосередитися на розробці адаптивних алгоритмів наступного покоління, що забезпечуватимуть автоматичну реконфігурацію структури комітетів у відповідь на зміни параметрів мережі. Такі алгоритми мають поєднувати функції виявлення аномалій, прогнозування навантаження та оптимізації розподілу ресурсів, враховуючи економічні стимули, різноманітні рівні продуктивності валідаторів та поведінкові особливості учасників при збереженні безперервності роботи механізму консенсусу.

Важливим кроком стане створення відкритого репозиторію стандартизованих датасетів та практичних реалізацій алгоритмів. Це полегшить стандартизацію експериментів, забезпечить високу повторюваність результатів та дозволить дослідникам зосередитися на розробці нових методів замість повторного відтворення відомих рішень.

Доцільно також розширити дослідження на інші PoS-протоколи (Cardano, Polkadot,

Solana) для оцінки універсальності підходу та його потенціалу масштабування в різних блокчейн-екосистемах. Це відкриває шлях до створення універсального фреймворку для оцінки алгоритмів розподілу валідаторів.

7. Висновки

Запропонований метод формування стандартизованих експериментальних наборів даних для порівняльної оцінки алгоритмів розподілу валідаторів у PoS-системах генерує сім ключових характеристик на основі статистичних параметрів Ethereum 2.0 з використанням фіксованих початкових значень. На відміну від існуючих досліджень, де параметри задаються менш формалізовано, запропоноване рішення забезпечує повну специфікацію всіх параметрів для детермінованої відтворюваності. Валідація підтвердила реалістичність методу через безпечну частку ненадійних валідаторів, незалежність економічних і технічних характеристик та відповідність мережевих затримок реальним показникам глобальних мереж.

Сформовано три стандартизовані набори даних різного масштабу з унікальними ідентифікаторами та метаданими. Датасети охоплюють сценарії від консорціумних до публічних мереж. Підхід мінімізує проблему несумісності експериментальних умов між дослідженнями. Статистичний аналіз згенерованих датасетів підтвердив логнормальний розподіл стейків, географічну децентралізацію між регіонами та відсутність кореляцій між результатами різних алгоритмів.

Розроблено систему комплексних критеріїв оцінювання: якість розподілу, стабільність результатів та обчислювальна ефективність. Система дозволяє оцінити компроміс між якістю оптимізації та обчислювальними витратами. На відміну від обмеженої оцінки в один або два виміри, вона забезпечує багатовимірний аналіз з урахуванням пропускної здатності, збалансованості навантаження та мережевих затримок.

Виконано системне порівняльне дослідження чотирьох алгоритмів на трьох масштабах з десятикратним повторенням. Виявлено масштабно-залежну поведінку: PSO+LS досягає найвищої якості оптимізації, але високі обчислювальні витрати обмежують його застосування періодичним офлайн-плануванням; EthShuffle забезпечує стабільні результати середньої якості; Random+Repair характеризується швидкістю виконання при варіативних результатах; Greedy демонструє максимальну швидкість та детерміністичність, однак його відносна ефективність знижується при малих масштабах мережі.

Проведено аналіз масштабованості, що виявив нелінійні залежності продуктивності від розміру мережі. На основі отриманих результатів сформульовано практичні рекомендації для відповідних сценаріїв застосування залежно від вимог до якості оптимізації та швидкості виконання. Запропоновані датасети та протоколи тестування створюють базовий інструментарій для об'єктивної оцінки нових алгоритмічних рішень у PoS-системах, забезпечуючи відтворюваність експериментів та можливість прямого порівняння результатів різних досліджень.

Перелік посилань:

1. E. Kapengut and B. Mizrach, “An Event Study of the Ethereum Transition to Proof-of-Stake,” *Commodities*, vol. 2, no. 2, pp. 96–110, Jun. 2023, doi: https://doi.org/10.3390/commodities2020006.

2. Ethereum Foundation, The Beacon Chain, 2024. URL: https://ethereum.org/en/upgrades/beacon-chain/.

3. Ethan “ethos.dev” (Joseph Ch), The Beacon Chain Ethereum 2.0 explainer you need to read first, ethos.dev, 2022. URL: https://ethos.dev/beacon-chain.

4. B. Edgington, Randomness, in: *Building Blocks*, Eth2 Book, 2025. URL: https://eth2book.info/latest/part2/building_blocks/randomness/.

5. T. Yan, S. Li, B. Kraner, L. Zhang, and C. J. Tessone, "A Data Engineering Framework for Ethereum Beacon Chain Rewards: From Data Collection to Decentralization Metrics," Scientific Data, vol. 12, no. 1, Mar. 2025, doi: https://doi.org/10.1038/s41597-025-04623-7.

6. Axon Trade, "Ethereum in 2025: Network, Usage, and Upgrades," Axon Trade, 2025. https://axon.trade/ethereum-in-2025 (accessed Aug. 25, 2025).

7. BlockByte, "The Fed's Pro-Crypto Pivot: How DeFi and Stablecoins Are Reshaping U.S. Financial Infrastructure," Ainvest, Aug. 21, 2025. https://www.ainvest.com/news/fed-pro-crypto-pivot-defi-stablecoins-reshaping-financial-infrastructure-2508 (accessed Aug. 25, 2025).

8. A. Leporati and L. Rovida, "Looking for Stability in Proof-of-Stake based Consensus Mechanisms," Blockchain Research and Applications, pp. 100222–100222, Jul. 2024, doi: https://doi.org/10.1016/j.bcra.2024.100222.

9. Q. H. Nguyen, A. Cronje, and M. Kong, "Fast Stochastic Peer Selection in Proof-of-Stake Protocols," arXiv (Cornell University), Jan. 2019, doi: https://doi.org/10.48550/arxiv.1911.04629.

10. V. T. Hoang, B. Morris, and P. Rogaway, "An Enciphering Scheme Based on a Card Shuffle," arXiv (Cornell University), Jan. 2012, doi: https://doi.org/10.48550/arxiv.1208.1176.

11. "Upgrading Ethereum | 2.9.4 Shuffling," Eth2book.info, 2025. https://eth2book.info/latest/part2/building_blocks/shuffling/ (accessed Aug. 27, 2025).

12. K. Venkatesan and S. B. Rahayu, "Blockchain security enhancement: an approach towards hybrid consensus algorithms and machine learning techniques," Scientific Reports, vol. 14, no. 1, p. 1149, Jan. 2024, doi: https://doi.org/10.1038/s41598-024-51578-7.

13. X. Gu, X. Wang, Y. Ma, Z. Chen, and S. Huang, "Performance Optimization for Information Sharing Process of BlockIoV Based on Multi-Objective Particle Swarm," pp. 172–183, Oct. 2023, doi: https://doi.org/10.1109/qrs60937.2023.00026.

14. H. Xiang, Z. Ren, Z. Zhou, N. Wang, and H. Jin, "AlphaBlock: An Evaluation Framework for Blockchain Consensus Protocols," arXiv (Cornell University), Jan. 2020, doi: https://doi.org/10.48550/arxiv.2007.13289.

15. "Towards a Unified Metric for Performance Evaluation of Proof-of-Stake Blockchains," Algorandtechnologies.com, 2018. https://algorandtechnologies.com/news/towards-a-unified-metric-for-performance-evaluation-of-proof-of-stake-blockchains (accessed Aug. 27, 2025).

16. W. Zhong et al., "Byzantine Fault-Tolerant Consensus Algorithms: A Survey," Electronics, vol. 12, no. 18, pp. 3801–3801, Sep. 2023, doi: https://doi.org/10.3390/electronics12183801.

17. "Metric: Validator Efficiency | beaconcha.in Knowledge Base," Beaconcha.in, Jun. 04, 2025. https://kb.beaconcha.in/v2beta/metric-validator-efficiency (accessed Aug. 28, 2025).

18. A. Lebedev and V. Gramoli, "On the Relevance of Blockchain Evaluations on Bare Metal," *arXiv (Cornell University)*, Jan. 2023, doi: https://doi.org/10.48550/arxiv.2311.09440.

19. "Deanonymizing Ethereum Validators: The P2P Network Has a Privacy Issue," Arxiv.org, 2024. https://arxiv.org/html/2409.04366v2 (accessed Aug. 28, 2025).

20. Y. Demenko, I. Grebennik, M. Kolmykov, Optimization of Validator Assignment to Committees in Proof-of-Stake Blockchain Systems, in: Modern Information Technologies and Artificial Intelligence Systems: Proceedings of the 1st International Scientific and Practical Conference MIT@AIS-2025, Kharkiv–Yaremche, May 19–22, 2025, KhNURE, Kharkiv, 2025, pp. 46–49. (In Ukrainian)

21. J. Kennedy, Particle Swarm Optimization, in: C. Sammut, G. I. Webb (Eds.), Encyclopedia of Machine Learning, Springer, Boston, MA, 2011. doi:10.1007/978-0-387-30164-8_630.

22. Y. Wang, "Review on greedy algorithm," Theoretical and natural science, vol. 14, no. 1, pp. 233–239, Nov. 2023, doi: https://doi.org/10.54254/2753-8818/14/20241041.

Надійшла до редколегії 08.09.2025 р.

Деменко Євгеній Євгенович, аспірант кафедри СТ ХНУРЕ, м. Харків, Україна, e-mail: yevhenii.demenko@nure.ua, ORCID: https://orcid.org/0000-0002-5699-1400

Гребеннік Ігор Валерійович, доктор технічних наук, професор, завідувач кафедри СТ ХНУРЕ, м. Харків, Україна, e-mail: igor.grebennik@nure.ua, ORCID: https://orcid.org/0000-0003-3716-9638

Колмиков Максим Миколайович, кандидат технічних наук, старший науковий співробітник, ХНУПС, м. Харків, Україна, e-mail: kolmykov.max@gmail.com, ORCID: https://orcid.org/0000-0003-1972-3543