

Децентралізована репутаційна мережа: Рамка для природної суспільної організації

Pavel Kudrna

Prague, Czechia

Чернетка v1 — березень 2026

Abstract

Почуття справедливості—переконання, що кривди виправляються, а заслуги винагороджуються—є найсильнішою згуртовуючою силою суспільства. Коли централізовані інституції врядування не можуть забезпечити це почуття, тому що вартість забезпечення права перевищує цінність самого права, суспільна згуртованість руйнується. Чинні інституційні структури не спроможні розв'язати значний клас спорів систематично так само, як централізоване планування не спроможне розподілити ресурси: через асиметрію інформації, відсутні зворотні зв'язки та відірваність тих, хто ухвалює рішення, від наслідків цих рішень. Ця стаття представляє репутаційну соціальну мережу (Reputation Social Network, RSN): децентралізований, непідкупний, стійкий до цензури комунікаційний рівень, побудований на децентралізованих ідентифікаторах (DID), що дає змогу псевдонімним учасникам публікувати, перевіряти та споживати репутаційну інформацію про реальну поведінку. Ключовий механізм спирається на три принципи проєктування: (1) асиметричну структуру вартості, за якої читання репутаційних даних дешево, але запис вимагає перевірюваного зусилля; (2) недетермінований протокол вибору верифікатора на основі консистентного хешування, який оцінює радикальні твердження через ескаляцію вартості ітерацій; та (3) ринкову модель репутаційного авторитету, у якій приватні верифікатори ставлять на кін накопичену репутацію на точність тверджень, які вони підтверджують. Отримана архітектура породжує емерджентну меритократію без центральної координації та уможливорює поступовий, оборотний шлях міграції від наявних систем, керованих державою.

1 Вступ

1.1 Проблема централізованої довіри

Сучасне врядування спирається на фундаментальне припущення: що централізовані інституції здатні опосередковувати довіру між незнайомцями в масштабі. Суди розв'язують спори. Реєстри засвідчують власність. Регуляторні органи забезпечують дотримання стандартів. Ці інституції були розроблені в епоху, коли інформація була дефіцитною, комунікація повільною, а делегування повноважень центральному органу було єдиним придатним механізмом координації. Проте централізоване врядування зазнає невдачі з тих самих структурних причин, що й централізоване планування: асиметрія інформації, відсутні зворотні зв'язки та відірваність тих, хто ухвалює рішення, від наслідків цих рішень.

Це припущення руйнується, наприклад, коли вартість інституційного посередництва перевищує цінність спору. Розгляньмо орендаря, який виявляє, що його орендованій квартирі бракує законодавчо необхідного сертифіката електробезпеки—стану, що становить безпосередню загрозу для життя. Законне право орендаря відмовитися від договору є однозначним. Проте вартість забезпечення цього права через судову систему перевищує грошову вартість застави та належного відшкодування, навіть з урахуванням можливої встановленої законом компенсації [1]. Рациональною відповіддю є змиритися зі втратою та вийти.

Це не патологічний винятковий випадок. Це стандартний досвід для значного класу спорів, у яких шкода реальна, але грошові ставки падають нижче порогу, за яким інституційне забезпечення стає економічно раціональним. Результатом є система, яка структурно сприяє недобросовісним акторам: ті, хто завдає шкоди іншим в обсягах нижче цього порогу, можуть діяти безкарно,

тому що вартість пошуку справедливості лягає на постраждалу сторону.

Наведений вище приклад узятو з правового середовища автора—чеської системи цивільного права. В інших правових традиціях—прецедентному загальному праві, звичаєвому праві, змішаних системах—справедливість зазнає невдачі по-різному і різною мірою, залежно від культурних та процесуальних особливостей юрисдикції. Читачі, найпевніше, упізнають еквівалентні приклади зі свого власного контексту. Структурно універсальним є сам патерн: спори, чия цінність падає нижче порогу економічно раціонального забезпечення, завершуються тим, що втрату поглинає постраждала сторона. RSN не обіцяє досконалої справедливості—вона лише зменшує структурну перевагу, якою користуються недобросовісні актори, коли інституційне забезпечення є неекономічним.

1.2 Чому наявні рішення не спрацьовують

Рамки децентралізованої ідентичності (DID) [2] надають криптографічні механізми для самосуверенного управління ідентичністю, але зосереджуються насамперед на верифікації ідентичності, не торкаючись ширшого питання поведінкової репутатії.

Прив'язані до душі токени (Soulbound Tokens, SBT) [3] відображають ідею про те, що репутатія непередавана, але покладаються на блокчейн-інфраструктуру з обмеженнями масштабованості і не розв'язують економічних стимулів, що керують внесенням інформації. Споріднені концепції, як-от токени заслуг (наприклад, Liberland), поділяють подібну філософію, але залишаються прив'язаними до конкретних юрисдикційних рамок.

Децентралізовані автономні організації (DAO) [4] реалізують врядування через смарт-контракти, але відтворюють плутократичну динаміку, де вплив пропорційний капіталу. Квадратичне голосування [5] частково пом'якшує це, але обмежує практичне впровадження. DAO також стикаються з фундаментальною *проблемою оракула*: смарт-контракти не можуть надійно перевіряти реальні стани без довіреного зовнішнього джерела, і ця проблема не має загального розв'язку в межах блокчейн-архітектури.

Жодна наявна система не поєднує децентралізацію, стійкість до цензури,

псевдонімність, перевірювану вартість входу та емерджентний консенсус.

1.3 Внесок

Ця стаття робить такий внесок:

1. Означення **репутаційної соціальної мережі (RSN)**, децентралізованого протоколу, що розширює рамку DID для підтримки двостороннього обміну репутатією.
2. **Недетермінований протокол вибору верифікатора** на основі консистентного хешування [6].
3. **Принцип дорогого радикалізму**, що оцінює твердження відповідно до їхньої віддаленості від консенсусу мережі через три сукупні канали вартості.
4. **Модель репутаційного авторитету** з асиметричними стимулами, де хибне підтвердження коштує катастрофічно більше, ніж легітимні збори.
5. **Поступовий шлях міграції** через паралельну фіскальну інфраструктуру.

2 Принципи проєктування

Архітектура RSN обмежена п'ятьма непорушними принципами проєктування.

Тяглість та еволюція. Система співіснує з наявними інституціями протягом перехідного періоду невизначеної тривалості. Перехід має бути оборотним на кожному етапі.

Добровільність та відповідальність. Участь є добровільною, але добровільність поєднана з відповідальністю: учасник, який бере на себе контроль над інституційною функцією, водночас бере на себе супутні зобов'язання.

Розмаїття та культурна нейтральність. Консенсус виникає з двосторонніх взаємодій, а не з наперед заданих правил, накиннутих проєктувальниками системи.

Стійкість та опір цензурі. Вартість цензурування мережі має перевищувати вартість її толерування. Лише повний тоталітаризм—за руйнівної політичної та економічної ціни—може придушити систему.

Консенсус та примус. Система включає людські схильності до дріб'язковості, злостивості та задоволення від відплати як несучі елементи. Неналежну поведінку не заборонено; її оцінено.

3 Огляд системи

3.1 Репутаційна соціальна мережа

RSN є децентралізованим одноранговим комунікаційним рівнем, у якому учасники обмінюються перевірюваною інформацією про реальну поведінку. Її визначальними властивостями є: децентралізована та нецензурована інфраструктура; псевдонімна участь через DID; асиметрична структура вартості (читання дешеве, запис дорогий); криптографічна перевірюваність; та **підтримка стандартів**— машиночитні формати для інформації, що поширюється мережею, для послуг, які пропонують авторитети (складання твердження, підтвердження, свідкова служба), та для формату політики, включно з правилами оцінювання репутації контрагента й оцінювання ризику невідповідності політики (між задекларованою та реальною поведінкою).

3.2 Архітектурні компоненти

RSN складається з чотирьох основних компонентів (рис. 1):

1. **Рівень DID.** Ідентичності учасників із задекларованими правилами, репутаційними метаданими та мережевою маршрутизацією.
2. **Протокол тверджень.** Структурований формат для публікації тверджень про реальні події.
3. **Мережа верифікації.** Децентралізований протокол призначення верифікаторів за допомогою консистентного хешування.
4. **Рівень агрегації репутації.** Служби, що виробляють зрозумілі людині зведення репутації.

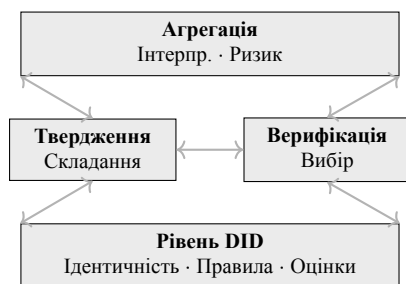


Figure 1: Чотирирівнева архітектура RSN.

3.3 Ролі учасників

Транзакція верифікації залучає до шести відмінних ролей DID (рис. 2): **Ініціатор** (DID_I , створює та подає твердження), **Суб'єкт** (DID_S , DID, про який твердження—може збігатися з Ініціатором), **Авторитет** (DID_A , підтверджує твердження за плату; може також пропонувати свідкові послуги—*опційно*), **Свідок** ($DID_{W_{1..n}}$, інкогніто-спостерігач за чесністю верифікатора через коди виклику—*опційно*), **Верифікатор** (DID_V , алгоритмічно обраний для публікації твердження) та **RSN** (мережа, де публікуються перевірені твердження). Будь-яку роль може бути делеговано іншому DID (розділ 7.4). Авторитет зазвичай поєднує підтвердження зі свідковими послугами, але ці дві функції логічно незалежні.

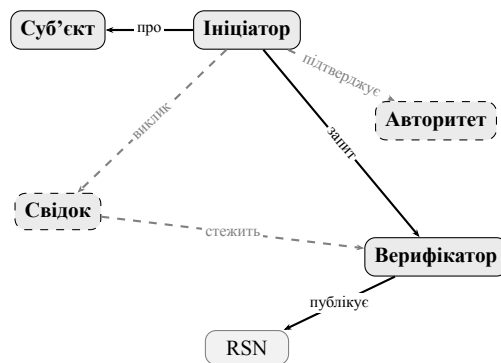


Figure 2: Ролі DID у транзакції верифікації. Пунктир = опційно (Авторитет, Свідок).

3.4 Непідкупність: чотири сили

Непідкупність RSN виникає не з єдиного механізму, а з чотирьох одночасних сил. Жодна не є достатньою сама по собі; лише їхнє поєднання робить спробу підкупу економічно ірраціональною.

Непередбачувана ціль. Верифікатор для кожного твердження обирається недетерміновано через консистентне хешування (розділ 5.3). Зловмисник не може заздалегідь націлитися на конкретного верифікатора для підкупу, тому що ідентичність верифікатора є функцією змісту твердження та попередніх ітерацій, а не вибору ініціатора.

Репутаційний важіль—повільно вгору, швидко вниз. Репутацію можна здобувати лише поступово, через тривалу послідовну поведінку. Проте її можна втратити катастрофічно в одному шахрайському підтвердженні (розділ 6.2). Ця асиметрія означає, що роки накопиченої

репутації можуть бути знищені одним нечесним підтвердженням; оптимальною стратегією залишається відхиляти підозрілі твердження.

Публічна обіцянка. Кожен Утримувач DID публічно декларує свою політику—машиночитний набір правил, яких він зобов’язується дотримуватися. Розходження між декларацією та реальною поведінкою є виявним і оцінюється як репутаційне порушення, тягче за саме базове порушення (розділ 7.3). Тому лицемірство коштує дорожче за пряму нечесність.

Асиметрія вартості атаки на сітку. Вартість цензурування чи дестабілізації мережі зростає нелінійно з розміром та щільністю мережі, тоді як вартість її толерування залишається сталою. Щоб цензура окупилася, централізованому актору довелося б застосувати її по всій мережі—що потребує заходів, непропорційних будь-якій мислимій вигоді (розділ 10).

4 Модель децентралізованої ідентичності

4.1 DID з особливими властивостями

RSN розширює специфікацію W3C DID Core [2] додаванням: **Задекларованих правил** (машиночитних поведінкових зобов’язань), **Репутаційних метаданих** (агрегованої поведінкової оцінки) та **Мережевої маршрутизації** (змінюваного опіон-шлюзу, окремого від стабільної позиції на кільці).

4.2 Життєвий цикл твердження

Твердження проходить через шість етапів (рис. 3): складання, опційне підтвердження авторитетом, вибір верифікатора через консистентне хешування, рішення верифікації (прийняти або відхилити з ітерацією), публікація та оновлення репутації для всіх сторін.

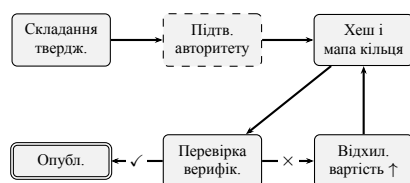


Figure 3: Життєвий цикл твердження з циклом ітерації.

4.3 Зв’язок із W3C DID Core

Модель ідентичності RSN є власною надмножиною специфікації W3C DID Core [2]. Усі DID у RSN є дійсними W3C DID. Специфічні для RSN розширення задовано як властивості DID-документа, означені у спеціальній специфікації методу DID, сумісній із наявною інфраструктурою, як-от ION [7] та Ceramic [8].

5 Верифікація та поширення інформації

5.1 Вартість внесення інформації

Ключовий економічний принцип RSN—асиметрія між читанням і записом. Читання репутаційних даних наближається до нульової граничної вартості для учасників, які є частиною мережі DID і мають доступ, співмірний з їхньою репутацією—новачки мусять поступово заслуговувати ширший доступ (див. антипаразиткування). Запис—що розуміється як тривала матеріалізація репутації в мережу, а не як одноразова технічна дія—вимагає перевірюваної сукупної інвестиції за трьома вимірами: **час** (роки життя, витрачені на послідовну поведінку, виконані зобов’язання та плекані стосунки), **енергія** (зусилля, вкладене в чесні справи, дбання про партнерства та довготривалу надійність) та **гроші** (інвестиція у власне ім’я—професійний розвиток, якість власної роботи, відшкодування завданих кривд). Репутацію не можна придбати миттєво—вона є результатом накопичення протягом усього життя, яке система лише робить видимим і передаваним між контекстами. Одноразові технічні витрати протоколу (криптографічні підписи, збори верифікаторів) мізерні порівняно з цією базовою інвестицією.

5.2 Соціальний граф та глибина контактів

RSN за своєю суттю є соціальною мережею: кожен DID додає контакти (інші DID), які дозволяють з’єднання. Ці контакти мають власні контакти, і так далі. Алгоритмічний пошук верифікатора діє в межах налаштовуваної глибини h пов’язаних контактів (наприклад, $h = 3$: власні прямі контакти, їхні контакти та контакти контактів). Ця архітектура не потребує єдиного глобального блокчейну—вона природно сприяє виникненню спільнот/кластерів із перекриттями

в інші спільноти. Кожен учасник DID мусить мати опубліковану **політику**—машиночитний набір правил, що керує тим, як оцінюються вхідні запити. Порушення політики записуються в мережу як негативні артефакти.

5.3 Алгоритмічний вибір верифікатора

Протокол верифікації використовує консистентне хешування [6] (рис. 4). Верифікація є платною послугою: верифікатори працюють за плату, створюючи ринок, де конкуренція формує ціни, швидкість та надійність.

Крок 1. Документ твердження конкатенується з результатом хешу попередньої ітерації (або $\emptyset$ на першій ітерації) та хешується:

$$\text{pos} = f_h(\text{claim} \parallel \text{prev_hash}) \quad (1)$$

Алгоритм мусить включати *повний шлях* усіх попередніх запитів та відповідей—а не лише хеш попередньої ітерації. Повна відповідь кожного верифікатора (прийняття, відхилення, названа ціна, причина) додається до документа, що зростає, і є перевірюваною через криптографічні підписи на кожному кроці.

Крок 2. Хеш відображається на N позицій на кільці консистентного хешу, рівномірно розподілених (наприклад, для $N = 4$: $+0^\circ, +90^\circ, +180^\circ, +270^\circ$). Від кожної позиції пошук за годинниковою стрілкою обирає найближчий DID як кандидата у верифікатори (рис. 5). N є змінним параметром, що може залежати від відсотка наразі активних DID, часу доби чи історичної відгукуваності мережі.

Крок 3. Верифікатор приймає (публікує, ставлячи на кін репутацію) або відхиляє (повертаючись до кроку 1 з хешем відхилення). Коли вузол не відповідає попри задекларований статус онлайн, наступний запит мусить включати всі відповіді від усіх N попередніх кандидатів у верифікатори.

Антипаразитизм. Цільові DID можуть встановлювати збори чи обмеження доступу для запитів від нових DID із малою репутацією. Цей градуїований механізм (не бінарний) змушує новачків будувати репутацію через справжню активність, перш ніж вільно споживати чужі дані.

Кожне відхилення додає повну відповідь верифікатора (включно з причинами та умовами) до документа твердження, розширюючи його зміст. З розширеного документа недетерміновано обчислюється новий хеш, що відображається

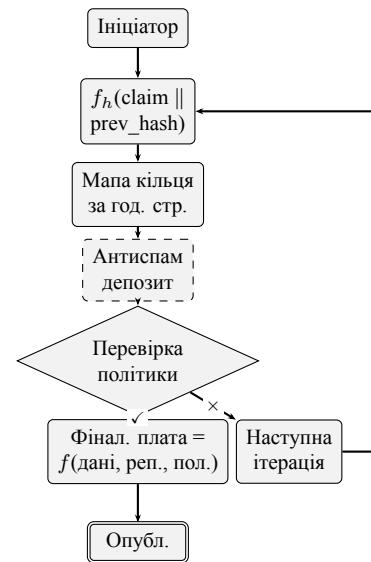


Figure 4: Протокол вибору верифікатора. Антиспам-депозит опційний і може бути поверненим. Фінальна плата сплачується лише за прийняття.

на іншу позицію кільця та обирає іншого верифікатора. Документування повного шляху є обов'язковим—ініціатор не може передбачити чи вплинути на наступного верифікатора. Якщо верифікатор ігнорує запит попри сумісну задекларовану політику, свідки (за наявності) фіксують це, і ініціатор може долучити свідкові докази при фінальній публікації.

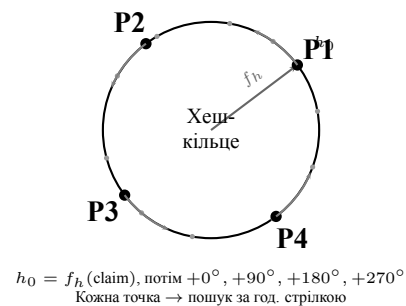


Figure 5: Багатоточковий вибір ($N=4$). Хеш h_0 задає зміщення; 4 точки рівномірно розподілені від h_0 .

5.4 Свідковий протокол

Роль **Свідка** забезпечує інкогніто-підзвітність за чесність верифікатора через криптографічний протокол кодів виклику:

Крок W1. Запитувач підписує запит на верифікацію та надсилає його N_w свідкам—контактам зі соціальної мережі запитувача або

спеціалізованим провайдером свідкових послуг, що працюють за плату.

Крок W2. Кожен свідок w_i підписує весь підписаний запит, зберігає оригінальний підпис σ_i та обчислює код виклику $c_i = h(\sigma_i)$. Код виклику додається до запиту.

Крок W3. Запит, що тепер несе N_w кодів виклику, надсилається верифікатору. Верифікатор бачить коди, але *не може визначити*, хто є свідками чи чи відповідають коди реальним підписам.

Крок W4 (чесний верифікатор). Якщо верифікатор відповідає згідно зі своєю задекларованою політикою, коди виклику залишаються непрозорими. Свідків ніколи не розкривають. Жодних додаткових даних не публікується.

Крок W5 (порушення політики). Якщо верифікатор порушує свою політику (ігнорує запит, відповідає непослідовно), запитувач утримує оригінальні підписи свідків σ_i . Їх можна опублікувати як супровідні дані: будь-хто може перевірити $c_i = h(\sigma_i)$, довівши, що свідок був присутній. Запитувач може опублікувати незалежно—верифікатор уже підписав коди виклику у своїй відповіді.

Властивість блефу. Запитувач може підставити випадкові значення замість деяких або всіх кодів виклику. Верифікатор не може відрізнити справжні коди (підкріплені високорепутаційними авторитетами) від шуму. Ця *невизначеність* є механізмом примусу: кожен запит несе ризик, що поважний авторитет спостерігає інкогніто. Вартість підтримання цього тиску майже нульова (генерація випадкових чисел), тоді як потенційна вартість нечесності для верифікатора катастрофічна. Чесна поведінка стимулюється навіть за відсутності реальних свідків.

Авторитет як свідок. Авторитет, що підтверджує твердження, може поєднати свідкові послуги: «Я підтверджую ваше твердження та служу інкогніто-свідком під час верифікації.» Це природна бізнес-модель—авторитет уже має контекст. Однак ці дві ролі логічно незалежні.

5.5 Принцип дорогого радикалізму

Три канали вартості накопичуються одночасно (рис. 6):

Канал 1: Вартість ітерації. Кожне відхилення змушує до нової ітерації, що споживає час і ресурси. Лінійне зростання.

Канал 2: Роздування документа. Кожна ітерація додає повну відповідь верифікатора до документа твердження. Зростання лінійне, але з базовим зміщенням: початкове навантаження (зміст твердження, підтвердження авторитету, криптографічні підписи) уже має нетривіальний розмір B_0 . Загальний розмір після k ітерацій: $S(k) = B_0 + k \cdot \Delta$, де Δ є середнім розміром відповіді.

Канал 3: Премія за ризик верифікатора. Ризик суб'єктивний. Верифікатор оцінює, чи задекларовані політики запитуваного DID суперечать власним комерційним, соціальним чи політичним інтересам верифікатора. Премія може зростати експоненційно зі сприйнятою віддаленістю від «ідеалу» верифікатора: $P(d) = \alpha \cdot e^{\beta d}$, де d є суб'єктивною метрикою відстані верифікатора. За персональною межею «нізачо» верифікатор може відмовити цілковито.

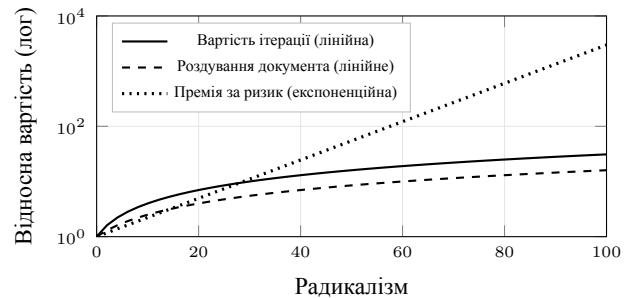


Figure 6: Ескалація вартості за трьома каналами. Премія за ризик домінує за високого радикалізму.

Критично, це не цензура. Жодне твердження не заборонене. Система оцінює ризик (табл. 1).

Table 1: Порівняння вартості за типами тверджень.

Тип	Ітер.	Док.	Плата	Разом
Достовірне	$k_{\min}$	B_0	$P_{\min}$	Низька
Спірне	k_{med}	$B_0 + k\Delta$	$\alpha e^{\beta d}$	Помірна
Радикальне	$k_{\max}$	$\gg B_0$	$\gg P_{\min}$	Дуже висока
Шахрайське	$\rightarrow \infty$	—	—	Неопублікована

6 Репутація та оцінювання ризику

6.1 Модель накопичення репутації

Репутація виявляє три властивості: **повільне накопичення** (поступове зростання через послідовну поведінку), **швидке руйнування** (одне шахрайство може катастрофічно знизити оцінку) та **індивідуальне оцінювання** (кожна

споживча сторона може застосовувати власну функцію агрегації).

6.2 Репутаційні авторитети

Репутаційний авторитет розглядає твердження і, якщо задоволений, співпідписує їх—ставлячи на кін власну репутацію (рис. 7). Асиметрія є задумом: здобуття репутації повільне (поступове $+\delta$ за чесне підтвердження), тоді як її втрата катастрофічна ($-\Delta \gg \delta$ за хибне підтвердження; ілюстративно, наприклад, $+2\%$ проти -70%). Оптимальною стратегією завжди є відхиляти підозрілі твердження. Конкретні значення δ та Δ є параметрами системи.

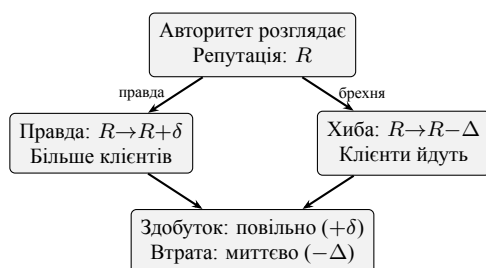


Figure 7: Репутаційний авторитет—асиметричні стимули.

6.3 Багатовимірна репутація

Репутація є не скаляром, а вектором за багатьма вимірами: фінансова надійність, особиста доброчесність, професійна компетентність, громадянська залученість та інші (рис. 8). Різні провайдери оцінювання проєктують цей вектор по-різному залежно від контексту—позикодавець надає пріоритет фінансовій надійності, роботодавець—професійній компетентності, сусід—громадянській поведінці. Немає єдиної «правильної» оцінки репутації; є лише перспективи.

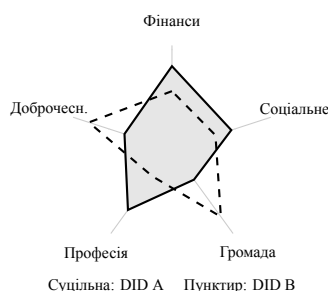


Figure 8: Багатовимірні вектори репутації. Різні DID виявляють різні профілі за вимірами.

6.4 Демократизоване оцінювання ризику

RSN демократизує систематичне оцінювання ризику—спроможність, наразі зосереджену у фінансових установах, але застосовну далеко за межами банківської справи. Промислові конгломерати, що оцінюють надійність постачальників, страхові компанії, що оцінюють поведінковий ризик, комплексна перевірка при злиттях і поглинаннях, оцінювання строку служби обладнання у виробництві—усюди, де ризик контрагента вимагає оцінювання, RSN надає децентралізовану, ринкову альтернативу. Ринок послуг інтерпретації перекладає сирі багатовимірні репутаційні дані на дієві зведення, роблячи оцінювання контрагента доступним будь-якому учаснику за граничну вартість.

7 Консенсус та розв'язання спорів

7.1 Модель децентралізованого правосуддя

RSN переформулює правосуддя як проблему двосторонніх перемовин. Загроза постійної, перевірюваної шкоди репутації є ефективнішим стримувавцем, ніж судові процеси, які постраждала сторона не може собі дозволити.

7.2 Емерджентний консенсус

Кожен учасник підтримує особистий поріг задоволення. Сукупний консенсус мережі виникає як статистичне середнє тисяч двосторонніх перемовин (рис. 9). Реакція, значно вища за консенсус (варварська), дорога для публікації. Реакція поблизу консенсусу (пропорційна) дешева. Консенсус є не правилом—він є цінним сигналом.

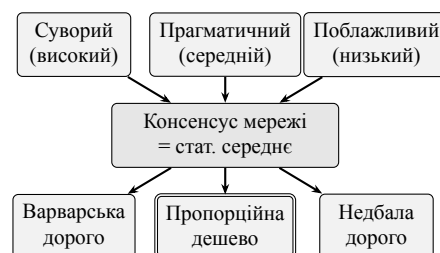


Figure 9: Емерджентний консенсус з індивідуальних порогів задоволення.

7.3 Калібрування покарання

Кожен Утримувач DID публічно декларує реакції на конкретні категорії неналежної поведінки. Непропорційно суворі чи поблажливі декларації притягують негативні репутаційні сигнали. Пропорційні декларації приймаються без тертя—самокалібрувальна система покарання.

Консенсусні декларації самі підлягають виявленню лицемірства: декларативне зобов'язання до консенсусної позиції за непослідовної поведінки становить репутаційне порушення, тягче за саме базове відхилення, оскільки демонструє навмисний обман.

7.4 Делегування

Усі дії в межах DID—за одним винятком—можуть бути делеговані іншому DID. **Роль Суб'єкта—DID, про чію поведінку йдеться у твердженні—не може бути делегована; вона неперерадано прив'язана до утримувача ідентичності, до якого стосується твердження.** Інші активні ролі—Ініціатор, Авторитет, Свідок, Верифікатор—можуть бути передані призначеному делегатному DID, чи то для верифікації, подання тверджень, участі в консенсусі чи розв'язання спорів. Делегування є відкличним у будь-який час. Це уможлиблює спеціалізацію (наприклад, професійні послуги верифікації), тоді як Утримувач зберігає остаточний контроль та відповідальність. Делегування застосовне по всій системі та є істотним для масштабованості.

7.5 Від індивідуальної моралі до емерджентного суспільного договору

Задекларована політика кожного DID являє собою формалізацію індивідуальної моралі: що Утримувач вважає прийнятним, як він реагує на конкретну поведінку, що вимагає від контрагентів. Ці політики не накинута проєктувальником системи—їх обирає кожен учасник і виражає в машиночитній формі.

Ця формалізація уможлиблює трьохетапну емерджентність. По-перше, індивідуальна мораль кодується як **політика**—набір задекларованих правил, порогів та функцій реакції, яких DID зобов'язується дотримуватися. По-друге, сукупність усіх політик у мережі виробляє **емерджентну етику**—статистично спостережувані норми, яких не спроєктував жоден окремих учасник, але які виникають

із взаємодії тисяч індивідуальних моральних позицій [9]. По-третє, ця емерджентна етика, виражена як спільні поведінкові норми, забезпечувані через двосторонні цінні сигнали, становить **вимірюваний суспільний договір**—не теоретичний конструкт, якого ніхто не підписував [10], а емпірично спостережуваний набір правил, підкріплений реальною поведінкою.

Цей процес віддзеркалює висновки теорії моральних основ [11]: людська мораль є інтуїтивною, плюралістичною та культурно змінною. RSN не вимагає морального консенсусу як входу; вона виробляє поведінковий консенсус як вихід. Отриманий суспільний договір не є статичним—він еволюціонує, коли учасники приєднуються, виходять і коригують свої політики у відповідь на зворотний зв'язок мережі. На відміну від класичної теорії суспільного договору, цей договір є відкличним, спостережуваним та забезпечуваним без суверена.

8 Економічна модель

Попередні розділи описали інструмент—механізми верифікації RSN, структуру вартості та емерджентний консенсус. Цей розділ описує методологію застосування цього інструмента до фіскального та управлінського переходу. Інструмент є універсальним; методологія нижче є одним із можливих застосувань.

8.1 Структура стимулів

Репутація як капітал створює прямі стимули до чесної поведінки. За нормальної роботи учасники будують репутацію природно через повсякденні транзакції та виконані зобов'язання. Основні витрати припадають на **негативні** твердження—фіксування шкоди, завданої іншими. Ця асиметрія стимулює корисну, надійну поведінку: бути чесним не коштує нічого зайвого, тоді як бути шкідливим створює задокументований слід, який інші платитимуть, щоб зберегти. Лицемірство—декларування правил без дотримання їх—є найдорожчим режимом провалу, оскільки демонструє навмисний обман.

8.2 Паралельна фіскальна система

Три компоненти уможлиблюють конкурентний тиск: (1) **Простий податок**—єдина пропорційна ставка без винятків, спочатку трохи нижче наявної

ефективної ставки; (2) **Електронна реєстрація витрат (ЕРВ)**—інверсія чеської моделі ЕЕТ, щоб громадяни наглядали за державними видатками; (3) **Спрямоване громадянами розподілення податків**—починаючи з кількох відсотків у перший рік і сягаючи, за десятиліття, будь-де від нижнього двозначного числа до повної міграції до системи, спрямованої громадянами, залежно від темпу впровадження. Реальний темп залежить від швидкості, з якою виникають вільноринкові альтернативи державним послугам, та від готовності держави співпрацювати з переходом; функція часу не мусить бути лінійною, і немає підстав встановлювати верхню межу того, куди впровадження може врешті прийти.

9 Шлях міграції

Міграція проходить через три фази (рис. 10): **Фаза 1: Накладення** (RSN як інформаційний рівень, держава—єдиний провайдер), **Фаза 2: Конкуренція** (RSN надає функціональні альтернативи, використання двох систем) та **Фаза 3: Перехід** (RSN перевершує державні еквіваленти, добровільна міграція). Перехід є оборотним на кожному етапі.



Figure 10: Трифазна міграція—оборотна на кожному етапі.

Основним механізмом тиску є фіскальний: коли умовні платники податків досягають «економічно значущої меншості X »—групи, достатньо великої, щоб репресії проти неї завдавали нетривіальної економічної шкоди, а громадянська непокоря ставала правдоподібною загрозою—держава вступає в перемовини. Для ілюстрації ми використовуємо $X \approx 15\%$ ВВП, але реальний поріг залежить від конкретної економіки.

10 Аналіз безпеки

Ми розглядаємо п'ять категорій супротивників: окремі недобросовісні актори, організовані групи, корпоративні супротивники, супротивники державного рівня та супротивники рівня протоколу.

Стійкість до Sybil [12]. Будувannya репутації вимагає тривалої, перевірюваної взаємодії. Вартість успішної атаки Sybil масштабується лінійно з кількістю фальшивих ідентичностей і квадратично з цільовим рівнем репутації. Керування кількома DID паралельно можливе, але дороге за задумом—кожна ідентичність мусить незалежно накопичувати репутацію через справжню активність. У вільних суспільствах ця вартість стримує безтурботне зловживання; у диктатурах паралельні DID стають інструментом виживання, що уможливило компартменталізований спротив, підпільну координацію та безпечнішу навігацію чорним ринком.

Захист від змови. Патерни взаємного підтвердження серед закритих груп є виявними через аналіз соціального графа. Функції агрегації репутації знецінюють твердження від тісно згрупованих джерел.

Супротивник державного рівня. Опіон-маршрутизація, відсутність централізованої інфраструктури та розподіл ролі Верификатора серед бази учасників забезпечують, що придушення вимагає заходів, непропорційних будь-якій вигоді.

Приватність проти підзвітності. Псевдонімність—золота середина між анонімністю та розкриттям ідентичності—дозволяє DID накопичувати значущу репутацію без розкриття реальної ідентичності Утримувача.

11 Міркування щодо приватності

Модель приватності RSN побудована на псевдонімності. Утримувач контролює розкриття ідентичності: мінімальне (чистий псевдонім), вибіркоче (пов'язані конкретні облікові дані) або повне (асоційована юридична ідентичність). Опубліковані твердження є постійними—система підтримує контекстуальну корекцію, але не стирання. Утримувач може відмовитися від скомпрометованого DID і почати заново, втративши накопичену репутацію.

12 Суміжні роботи

Специфікація W3C DID Core [2] та мережа Ceramic [8] надають основу ідентичності. EigenTrust [13] продемонстрував розподілену агрегацію репутації без центрального авторитету. SBT [3]

запровадили непередавані соціальні токени. RSN відрізняється акцентом на економічній вартості як фільтрі якості та своїм механізмом оцінювання радикалізму.

DAO [4] та квадратичне голосування [5] продемонстрували здійсненність децентралізованого врядування. RSN виводить консенсус із двосторонніх цінових перемовин, а не з голосування.

Пропозиція спирається на анархо-капіталістичну теорію [14, 15] щодо приватного надання послуг, але відходить у двох критичних аспектах: вона проектує під злостивість та імпульси відплати, а не припускає раціональність, і пропонує поступовий перехід, а не революцію. Концепція емерджентних суспільних договорів має попередників у теорії спонтанного порядку Гаєка [16].

Анархо-агоризм. RSN поділяє значну спільну основу з агористською контр-економікою [17]—зокрема акцент на побудові паралельних інституцій та добровільному обміні. Однак агоризм схильний припускати раціональний власний інтерес як достатній механізм координації і часто бракує конкретного шляху міграції від наявних державних структур. RSN явно включає нераціональні поведінкові схильності та надає механізм фіскального тиску для поступового переходу.

Меритократія. RSN може являти собою перший функціональний опис того, як меритократія [18] може виникнути як системна властивість, а не як гасло. Традиційні меритократичні системи зазнають невдачі, тому що потребують центрального авторитету для означення та забезпечення «заслуги». У RSN заслуга виникає з двосторонніх оцінювань: ті, хто постачає цінність, накопичують репутацію; жоден комітет не вирішує, хто має заслугу.

Власність як привілей. RSN фундаментально відходить від анархо-капіталістичного та австрійсько-школового трактування прав власності як природних та абсолютних. У рамці RSN власність є *привілеєм*—суспільним визнанням, яке в крайніх випадках може бути відкликane спільнотою, коли учасник фундаментально порушує спільні норми (зрада, відмова захищати спільноту в екзистенційній кризі, систематична експлуатація). Це не державна експропріація, а відкликання суспільного визнання мережею двосторонніх стосунків.

13 Обговорення та обмеження

Холодний старт. Цінність RSN залежить від мережевих ефектів; ранні адоптанти стикаються з обмеженою цінністю, доки участь не досягне порогу. Однак навіть із ранніх стадій мережа DID слугує корисним допоміжним джерелом інформації. Очікується, що раннє оцінювання репутації та оцінювання авторитетів розвиватимуться поступово з наростанням витонченості.

Антипаразиткування та контроль доступу. Цільові DID можуть накладати градуйовані збори та обмеження доступу на запити від низькорепутаційних DID. Це не бінарне, а неперервна шкала: що менше репутації має запитуваний DID, то менше інформації він отримує, то довше чекає й то більше платить. Цей механізм стимулює справжню участь понад пасивне споживання.

Нерівність доступу. Вартість публікації тверджень може відфільтровувати легітимні скарги економічно знедолених учасників. Пом'якшення: кожен учасник водночас слугує потенційним верифікатором (або делегує цю роль) і заробляє збори за верифікацію. За достатнього часового горизонту нерадикальний учасник має наблизитися до фінансової нейтральності—дохід від верифікації приблизно компенсує витрати на публікацію. Це статистичне очікування, а не гарантія.

Нерівність репутації. Ранні адоптанти накопичують переваги, що можуть створити бар'єри для пізніших учасників. Однак оцінювання репутації виконується сторонніми провайдерами, які можуть обрати пом'якшення переваги першопрохідця через свої алгоритми агрегації. Бути першим із хорошою репутацією є легітимною порівняльною економічною перевагою—і це за задумом.

Відкриті питання. Оптимальні функції вартості, стандарти агрегації, врядування протоколом та взаємодія зі синтетичними репутаційними даними, згенерованими ШІ, залишаються сферами майбутньої роботи.

Ця стаття не стверджує, що RSN вироблятиме оптимальні результати за всіх обставин. Вона стверджує, що RSN являє собою *здійсненню* альтернативу—технічно реалізовану, економічно самопідтримувану та соціально сумісну з людськими поведінковими схильностями такими, якими вони насправді є.

14 Висновок

Ця стаття представила репутаційну соціальну мережу, децентралізований протокол, що уможлиблює псевдонімний, перевірюваний обмін репутацією. Принцип дорогого радикалізму забезпечує, що шахрайські твердження виключаються ціною без цензури. Запропонований шлях міграції уможлиблює поступовий, оборотний перехід від наявних інституцій. Проект включає людську природу такою, якою вона є—разом із дріб'язковістю, злостивістю та бажанням задоволення від відплати—замість того, щоб вимагати ідеологічної трансформації. Результатом є не утопія, а система, де правосуддя доступне, репутація заслужена, а вартість неналежної поведінки несе сторона, яка її спричинила.

References

- [1] Czech Republic, “Act no. 549/1991 coll., on court fees, as amended,” 1991, fee schedule for court proceedings; cost rules per Act No. 99/1963 Coll. (Code of Civil Procedure) and Decree No. 177/1996 Coll. (Attorney Tariff).
- [2] M. Sporny, D. Longley, M. Sabadello, D. Reed, O. Steele, and C. Allen, “Decentralized identifiers (DIDs) v1.0,” W3C Recommendation, Jul. 2022.
- [3] E. G. Weyl, P. Ohlhaver, and V. Buterin, “Decentralized society: Finding Web3’s soul,” *SSRN Electronic Journal*, May 2022.
- [4] V. Buterin, “DAOs, DACs, DAs and more: An incomplete terminology guide,” Ethereum Blog, May 2014.
- [5] V. Buterin, Z. Hitzig, and E. G. Weyl, “A flexible design for funding public goods,” *Management Science*, vol. 65, no. 11, pp. 5171–5187, 2019.
- [6] D. Karger, E. Lehman, T. Leighton, R. Panigrahy, M. Levine, and D. Lewin, “Consistent hashing and random trees: Distributed caching protocols for relieving hot spots on the World Wide Web,” in *Proc. 29th ACM STOC*, 1997, pp. 654–663.
- [7] D. Buchner, “ION — a scalable DID method based on Bitcoin,” Microsoft, 2021.
- [8] Ceramic Network, “Ceramic protocol specification,” GitHub, 2023. [Online]. Available: <https://github.com/ceramicnetwork>
- [9] É. Durkheim, *The Division of Labor in Society*. Free Press, 1893, english translation 1984.
- [10] J. Rawls, *A Theory of Justice*. Harvard University Press, 1971.
- [11] J. Haidt, *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. Vintage Books, 2012.
- [12] J. R. Douceur, “The Sybil attack,” in *Proc. 1st International Workshop on Peer-to-Peer Systems (IPTPS)*, 2002, pp. 251–260.
- [13] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina, “The EigenTrust algorithm for reputation management in P2P networks,” in *Proc. 12th International Conference on World Wide Web*, 2003, pp. 640–651.
- [14] M. N. Rothbard, *For a New Liberty: The Libertarian Manifesto*. Macmillan, 1973.
- [15] H.-H. Hoppe, *Democracy: The God That Failed*. Transaction Publishers, 2001.
- [16] F. A. Hayek, *Law, Legislation and Liberty*. University of Chicago Press, 1973.
- [17] S. E. I. Konkin, *An Agorist Primer*. KoPubCo, 2006.
- [18] M. Young, *The Rise of the Meritocracy*. Thames and Hudson, 1958.