

В. С. ВОЛОШИН

доктор технічних наук, професор,
професор кафедри промислової теплоенергетики
Приазовський державний технічний університет
ORCID: 0009-0005-6809-6779

ЩОДО ПИТАННЯ ПРО ОЦІНКУ ЯКОСТІ ІНФОРМАЦІЇ В ІНТЕРНЕТІ БЕЗ ЗАЛУЧЕННЯ ЦЕНЗУРИ

Робота присвячена проблемам розробки методів спрощеного аналізу достовірності інформації в мережі інтернет, як мети дослідження. Пропонується звернути увагу на такий об'єктивний показник інформації, як її ентропія, міра невизначеності або випадковості повідомлення. Показано, що сама по собі інформаційна ентропія не визначає, правдива чи хибна інформація належить аналізу. Однак вона здатна відображати якість інформації, пов'язану зі структурною визначеністю і передбачуваністю, що може бути використано як ознака правдоподібності. Ентропія поєднується зі структурною зв'язністю інформації. Це не абсолютний критерій істини, але зміни в структурній когерентності можуть бути показником потенційного обману, особливо якщо порівнювати з іншими структурами повідомлення. Приклади виявлення у складі текстового повідомлення таких критеріїв визначеності, як когерентність, банальність і хаотичність, за допомогою яких можна не тільки структурувати текстові матеріали, але й виділяти в них такі складові, які показують на достовірність цих матеріалів. Наведено приклади вимірювання інформаційної ентропії Шеннона для текстових повідомлень різної конфігурації, за допомогою якої можна співвимірювати інформацію за означеними вище функціями. Створено алгоритм пошуку неправдивої інформації в текстових повідомленнях, за допомогою якого можна послідовно виявляти режими об'єктивності сукупних інформаційних текстових структур в межах прийнятих параметрів – функції когерентності, банальності, хаотичності. Запропоновані співвідношення цих функцій в інформаційному просторі «правда-фальшивість». Запропоновано чек-лист із найбільш коректних і об'єктивних відношень між зазначеними в повідомленні параметрами, який дозволяє надавати формалізовану оцінку рівнів довіри текстового повідомлення. Для оцінки стану формалізованих повідомлень запропоновано для використання радар-графіки у сумісних координатах: ентропія-когерентність-банальність-хаотичність для певної групи слів, що міститься у повідомленні. Запропонована методика оцінки рівнів неправдивості текстових повідомлень за прямими і непрямими показниками, незважаючи на її формалізацію, дозволяє в першому наближенні оцінити достовірність конкретних повідомлень, більш об'єктивно підходити до інформації в мережі Інтернет, яка вже довгий час обходиться без критики або цензури, і має стійку тенденцію до розвитку фальшивості цілих інформаційних потоків.

Ключові слова: фальшива інформація, неправда-правда, інформаційна ентропія, когерентність, банальність, хаотичність, повідомлення, структура повідомлення.

V. S. VOLOSHYN

Doctor of Technical Sciences, Professor,
Professor at the Department Energy Sciences
Priazovsky State Technical University
ORCID: 0009-0005-6809-6779

ON THE ISSUE OF ASSESSING THE QUALITY OF INFORMATION ON THE INTERNET WITHOUT THE INVOLVEMENT OF CENSORSHIP

The work is devoted to the problems of developing methods for simplified analysis of the reliability of information on the Internet, as the goal of the study. It is proposed to pay attention to such an objective indicator of information as its entropy, the measure of uncertainty or randomness of the message. It is shown that information entropy itself does not determine whether the information is true or false for analysis. However, it is able to reflect the quality of information associated with structural certainty and predictability, which can be used as a sign of plausibility. Entropy is combined with the structural coherence of information. This is not an absolute criterion of truth, but changes in structural coherence can be an indicator of potential deception, especially when compared with other message structures. Examples of identifying such criteria of certainty as coherence, banality and chaotic in a text message, with the help of which it is possible not only to structure textual materials, but also to highlight such components in them that indicate the reliability of these materials. Examples of measuring Shannon's information entropy for text messages of different configurations are given, with the help of which it is possible to co-measure information according to the above functions. An algorithm for searching for false information in text messages has been created, with the help of which it is possible to consistently identify the modes

of objectivity of aggregate informational text structures within the accepted parameters – the functions of coherence, banality, chaoticness. The correlations of these functions in the truth-falsity information space are proposed. A checklist of the most correct and objective relations between the parameters specified in the message is proposed, which allows providing a formalized assessment of the confidence levels of a text message. To assess the state of formalized messages, it is proposed to use radar graphics in compatible coordinates: entropy-coherence-banality-chaotic for a certain group of words that contained in the message. The proposed methodology for assessing the levels of falsity of text messages by direct and indirect indicators, despite its formalization, allows in the first approximation to assess the reliability of specific messages, to more objectively approach information on the Internet, which has been doing without criticism or censorship for a long time, and has a steady tendency to develop falsity of entire information flows.

Key words: false information, false-truth, informational entropy, coherence, banality, chaotic, message, message structure.

Постановка проблеми

Очевидні обсяги недостовірної, дубльованої, неправдивої інформації, якою насичений сучасний Інтернет, і яка знаходиться у відкритому користуванні сотень мільйонів людей не тільки з числа звичайних читачів, але і з числа фахівців – фінансистів, вчених, політиків, для яких достовірність інформації є об'єктивною потребою. В умовах відсутності контролю за достовірністю такої інформації, здатністю неправди і брехні «маскуватися» під достовірність, шкода, що наноситься нею, вже порівнюється зі шкодою, завданою економічними, екологічними та іншими глобальними проблемами. Тому пошук методи виявлення такої неправдивості актуальні не тільки для наукового інформаційного простору, а й для суспільства в цілому.

Аналіз останніх досліджень і публікацій

За різними оцінками, інформація, що розміщується в інтернеті до 2025 року, досягла 175–200 ZB, збільшившись за останні п'ять років більш ніж у чотири рази (дані IDC-International Data Corporation). Структура цієї інформації, її спрямованість і способи впливу на сучасне суспільство стають все більш цікавими. Не останнє місце в таких дослідженнях займає об'єктивність і достовірність даної інформації.

Первинна інформація, вихідні дані, створені безпосередньо джерелом (наукові статті, офіційні документи, первинний користувацький контент) за даними [1], складають 10–20 % від загального обсягу (~14,4–28,8 ZB). Вторинна інформація, повтори, копії, інтерпретації, компіляції первинних даних (репости, рерайти, переклади) складають ~70 % інтернет-трафіка з 3 млн акаунтів та 4,5 млн твітері [2]. А за даними [3], це навіть 80–90 % (~115,2–129,6 ZB) всієї інформації в інтернеті. У будь-якому випадку в інтернеті переважає вторинна інформація.

Крім вторинної інформації, аналізу підлягає недостовірна інформація та інформація неправдивого характеру. Різні дослідження (Каліфорнійський університет, платформа DemandSage, міжнародна компанія GBTIMES, Фінляндія) показують що неправдива або хибна інформація може становити від 47,3 % до майже 79 % від сукупної інформації, що розташована в інтернеті, але точність залежить від методології дослідження. Таке розпорошення вже не є важливим, якщо врахувати, що дезінформація й у значно менших кількостях впливає на свідомість суспільства, якість наукових знань, об'єктивність прийнятих рішень, у тому числі на найвищому рівні сучасної політики або соціуму. Така інформація сприяє великій кількості різнооб'ємних кризових явищ в суспільстві, в науці.

Роботи С. Чжана та його колег [4] у галузі порівняльного моделювання для виявлення фейків за допомогою чисельних метрик та інших методів аналітики, а також методи К. Шу, Х. Лю та ін. [5, 6] у галузі моделей раннього виявлення на орієнтованих графах, надають можливостей для зниження ризику дезінформації та її оцінки у порівнянні з бенчмарком. Однак в такій же мірі, як і методи, описані в роботах [7, 8] по сукупній оцінці довіри до знань в алгоритмах ранжування інформації, вони дещо перевантажені критичними параметрами, різноманіття яких ставить ці методи в число чисто наукових і не застосовних на практиці. Уміння розпізнавати неправдиву інформацію різноманітними способами – аналітичними, експертними, топологічними [9, 10], за умовами їх нескладності для користувача, означає здатність уникнути величезної кількості помилок, нервових потрясінь, банкрутств, політичних втрат. А здатність запобігати та усувати спотворену та неправдиву інформацію може бути не менш важливою, ніж забруднення навколишнього середовища.

Формулювання мети дослідження

Створення простого за використанням і ефективного методу тестування інформації на предмет виявлення в ній недостовірних або неправдивих фрагментів та текстів.

Викладення основного матеріалу дослідження

У роботі пропонується звернути увагу на такий об'єктивний показник інформації, як її ентропія. Це міра невідзначеності або випадковості повідомлення. Сама по собі ентропія не визначає, правдива чи хибна інформація. Однак це впливає на якість інформації, пов'язаної зі структурною визначеністю і передбачуваністю, що може бути використано як ознака ступеня правдоподібності [11].

1. Правдива інформація пов'язана з помірною ентропією, оскільки містить ієрархічно взаємопов'язане і структуроване, логічне і зв'язне подання даних, передбачувані зв'язки між частинами на основі знань про реальний світ.

2. Неправдива інформація може мати аномально високу або аномально низьку ентропію, що характеризується великою кількістю неструктурованих, незрозумілих і суперечливих фрагментів. Якщо ентропія близька до максимуму, то елементи однаково ймовірні, а повідомлення близьке до «шуму», до хаосу. Якщо ентропія близька до нуля, повідомлення є передбачуваним до точки примітивності.

Ентропія поєднується зі структурною зв'язністю інформації [12, 13]. Це не абсолютний критерій істини, але зміни в структурній когерентності можуть бути показником потенційного обману, особливо якщо порівнювати з іншими структурними особливостями повідомлення.

Під структурною когерентністю інформації ми розуміємо ступінь зв'язності, послідовності і логічності елементів повідомлення, яка виражається через впорядкованість розподілу інформації і зменшення несподіванок між частинами тексту. З точки зору інформаційної ентропії Шеннона, структурна когерентність – це нерівномірний, неусвідомлений розподіл ймовірностей появи елементів, де кожне наступне повідомлення передбачувано в контексті попереднього, але не тривіально. Це підтверджується об'єктивністю сучасних систем штучного інтелекту, де лінгвістичні екстраполяції вербальних структур є головною об'єктивністю генерування нових когерентних текстів. Когерентний текст тут розглядається як не надто хаотичний і не надто шаблонний. – він оптимально організований для передачі сенсу. Локальна ентропія в ньому (на рівні словосполучень, абзаців, окремих слів) різко не скаче і має закономірності, що відображають логіку і смислову зв'язність.

Розглянемо приклад зв'язного, помірно ентропійного, логічного повідомлення на кшталт: *«Економічне зростання країни збільшилося в другому кварталі, чому сприяло зростання внутрішнього споживання і стабільність на світових ринках»*. Тут важливі такі параметри:

- передбачуваність, коли кожна наступна фраза логічно впливає з попередньої;
- помірна ентропія, що заснована на виборі строго обмежених контекстом слів;
- висока когерентність, що пов'язано з причинно-наслідковими зв'язками між словосполученнями.

Припустимо, що $H_k(p)$ ентропія Шеннона для зв'язного висловлювання. Тоді $N(H)$ – функція свідомої зв'язковості (когерентності) що визначається як $N(H) = -|H - H_{opt}| + \text{const}$, де $H_{opt} = \log_2(p)$ – оптимальна (когерентна) ентропія явно нетривіального тексту.

У зв'язному тексті, про який йде мова, всього 17 слів. Всі слова унікальні, ймовірність кожного $p = 1/17 = 0,0588$. Ентропія кожного слова $H_k(p) = -p \cdot \log_2(p) = -0,0588 \cdot \log_2(0,0588) = 0,2404$. Суммарна ентропія повідомлення $H = 17 \cdot 0,2404 = 4,0868$.

І тоді ми наведемо приклад навмисно незв'язного, високоентропійного або хаотичного повідомлення: *«Розвиток переривається, автомобіль рухається, розвинена автомагістраль, розрив, розвиток, дорога, кін, автомобіль»*. Він має хаотичну структуру з низькою передбачуваністю, відсутністю когерентності та логічних зв'язків у тексті. Для нього показник функції хаотичності визначається як $C(H) = -H \cdot D$, де D – міра дисперсії розподілу хаосу.

У цьому хаотичному тексті всього 11 слів. Два слова повторюються, а 9 слів є унікальними. Для них $p_1 = \frac{2}{11} = 0,1818$ і $p_i = \frac{1}{11} = 0,0909$, а $\bar{p} = \left(\frac{1}{10}\right) \cdot \left(\frac{2}{11} + 9 \cdot \frac{1}{11}\right) = 0,1$. Дисперсія $D = \frac{1}{n} \sum (p_i - \bar{p})^2 = 0,0069$ є досить високою для такої виборки. Сумарна ентропія заданого хаотичного виразу $H_h^{11} = -[0,1818 \cdot (-2,459) + 9 \cdot 0,0909 \times (-3,468)] = 3,2815$. Функція хаосу в такому повідомленні $C(H) = -3,2815 \cdot 0,0069 \approx -0,0226$. Досить низьке і негативне значення вказує на низьку хаотичність в моделі повідомлення. Тобто розподіл слабко нерівномірний, та інформаційна ентропія низька. Якщо б у виразі було більше нерівномірних частот (наприклад, одне слово повторювалося 5 разів, решта повторювалися по одному разу), а ентропія залишалася високою, то випадковість $C(H)$ була б більш вираженою ($\sim -0,0554$).

І проміжний приклад некогерентного, але низькоентропійного, стереотипного повідомлення: *«Китай зростає. Китай сильний. Китай завжди перемагає. Китай – найкращий»*. Воно, перш за все, банальне (нова функція), про що свідчить надмірна передбачуваність, низька ентропія і, особливо, формалізована, але не змістовна зв'язність, що виражається в повторенні банальних гасел.

Тут функція банальності зменшується зі зростанням H в експоненціальній залежності: $-B(H) = \exp(-\lambda H)$. Банальний текст містить усього 9 слів: «Китай» – 4 рази: $p = 4/9 = 0,4444$; $H_e^4 = -p \cdot \log_2(p) = 0,4444 \cdot \log_2(0,4444) = -0,52$. Решта п'ять слів, для кожного $p = 1/9 = 0,1111$ $H_e^1 = -p \cdot \log_2(p) = 0,3522$. Ентропія кожного слова $H_e^1 = -p \cdot \log_2(p) = -0,0909 \cdot \log_2(0,0909) = 0,3145$. Суммарна ентропія $H_e^9 = H_e^4 + 5H_e^1 = 0,52 + 5 \cdot 0,3522 = 2,281$.

Отже, розрахункова ентропія для порівнянних текстів: когерентного – $H = 4,0868$; хаотичного – $H = 3,2815$; банального – $H = 2,281$. Це свідчить про те, що семантична новизна (когерентність) – це максимум інформативності при структурованому її викладі. Ентропія висока, але не максимальна. Хаотичний посил – це неорганізованість, випадковість. Ентропія може бути високою, але сенс не читається. А банальний меседж – це надмірне повторення, гасла. Ентропія низька, тому що елементи передбачувані і однорідні.

Таким чином, можна припустити, що фальшивість текстового повідомлення часто знаходиться в зонах надмірної випадковості (високий H і високий D) або банальності (низький H і високий рівень повторень). Істина

може існувати в зоні структурованої новизни ($H \approx 3,8$ і $D \approx 1 \cdot 10^{-4}$). Відповідно, будемо виділяти три режими інформаційної структури (табл. 1).

Інтерпретація інформаційних категорій, що нас цікавить в контексті інформаційної фальші, має свої відмінності. Зокрема, банальні тексти, шаблонна інформація мають низьку ентропію і майже не несуть новизни. Вони легко запам'ятовуються, піддаються натякам, тиражуванню і можуть маскувати пропаганду. Хаотичні тексти створюють відчуття складності, «інформаційного шуму». Це може бути методом приховування дезінформації або ускладнення її верифікації. Когерентні тексти ближче до правди, вони мають впорядковану структуру, помірну передбачуваність, схильні до структурованої невизначеності.

Таблиця 1

Режими інформаційних структур для повідомлень в межах прийнятих параметрів

Категорія інформації	Ентропія	Структура	Зв'язність (когерентність)	Змістова новизна	Типова неправда
Банальна	Низька	Повторювана, шаблонна	Висока формально, але низька по суті	Низька	Пропаганда, кліше
Хаотична	Висока	Випадкова, фрагментарна	Низька	Низька	Фальсифікація, дезінформація
Когерентна	Середня	Логічно впорядкована	Висока	Висока	Маловірогідна, частіше правда

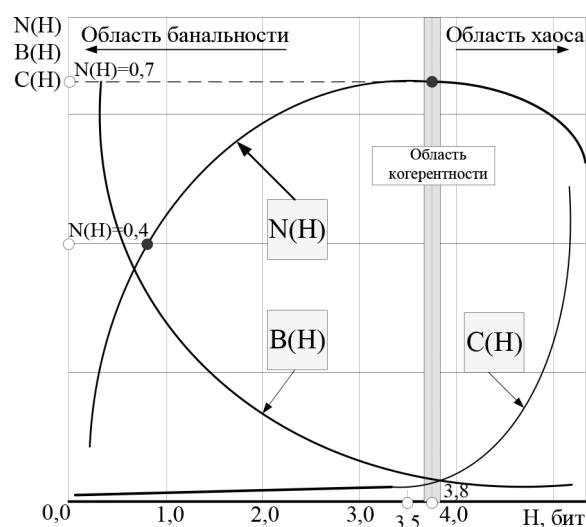


Рис. 1. Співвідношення когерентного, банального і хаотичного в інформаційному просторі «правди-неправди». Тут: $N(H)$ – функція когерентності; $B(H)$ – функція банальності; $C(H)$ – функція хаотичності

Неважко помітити, що на досліджуваних залежностях (рис. 1) є особливі точки, а саме, при $H < 3,5$ випадковість інформації майже дорівнює нулю, а при $H > 3,5$ випадковість інформації починає різко зростати. Це означає, що висока ентропія в поєднанні з відсутністю структурованості і високою дисперсією робить текст хаотичним, важким для читання і підозрілим на об'єктивність.

Вертикальна лінія, що проходить через $H = 3,8$ ж це область максимальної новизни інформації, коли її банальність майже зникає, а хаос ще не настав. Такий діапазон значень ентропії може бути характерний для об'єктивно інформативних, логічних, когерентних тверджень.

Безумовна теорія відносності в цьому випадку присутня.

Запропонований алгоритм пошуку та відрізнєння неправдивих текстових повідомлень в мережі Інтернет за параметрами, що розглядаються, включає наступне (рис. 2).

Збір статистичних даних по текстовому повідомленню, з розбивкою тексту на окремі слова, словосполучення за змістовним значенням з визначенням частоти слів в угрупованнях. На основі цієї інформації про повідомлення розраховується ентропія структури тексту, оцінюються параметри когерентності $N(H)$, банальності $B(H)$, хаотичності $C(H)$ тексту в порівнянні з їх ентропією. Такий алгоритм дозволяє аналізувати текстові повідомлення в мережі Інтернет на предмет їх достовірності в залежності від структурної ентропії тексту, використовуючи поняття когерентності, банальності і хаотичності інформації в повідомленні. Модель може бути представлена у вигляді Python-скрипта для конкретних рішень, що дозволяє розрахувати ентропію Шеннона, когерентність, банальність і хаотичність на основі статистики повторення слів у тексті, і зробити формалізовані висновки про його достовірність.

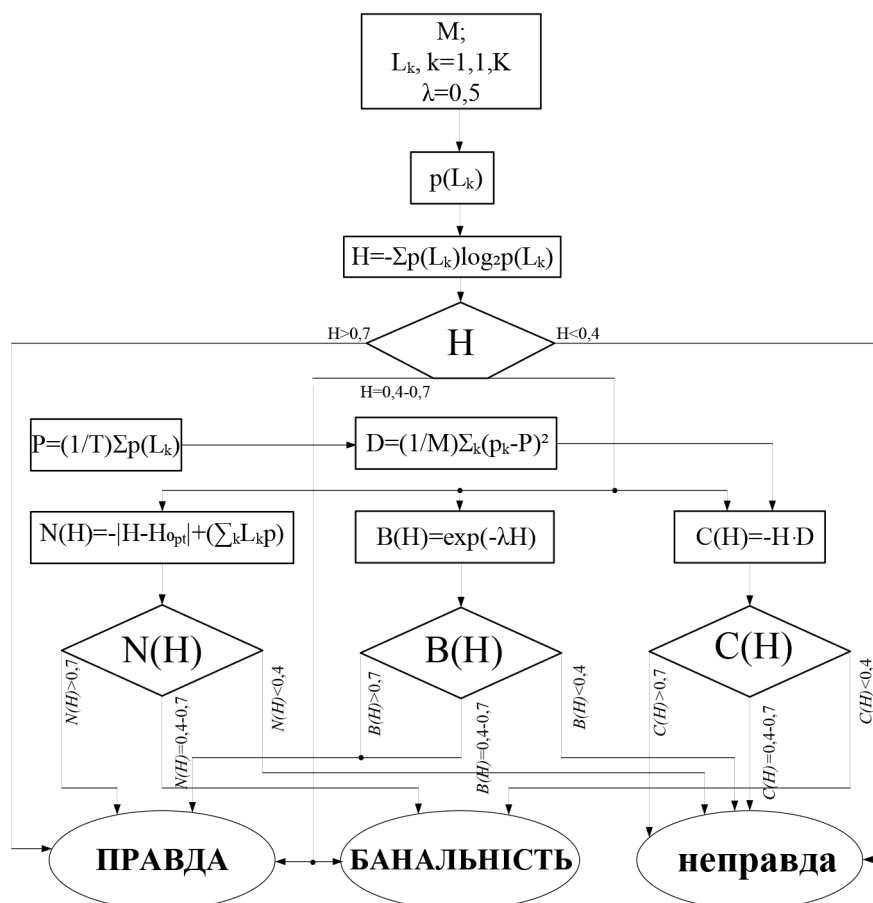


Рис. 2. Блок-схема пошуку неправдивої інформації в текстових повідомленнях.

Тут: M – загальна кількість слів у повідомленні; L_k – кількість слів у k -й групі однотипних слів; λ – коефіцієнт банальності виразу; $p(L_k)$ – частотність слова в k -й групі однотипних слів

Зі структури самого алгоритму видно, що під найбільшим числом варіантів розподілу ентропії мають бути на увазі хаотичні стани інформації з відносно невеликою ентропією, що знаходяться в області інформаційної рівноваги.

Для аналізу достовірності тексту повідомлення, що підлягає дослідженню на достовірність, сформуємо чек-лист з найбільш коректних і об'єктивних відношень між зазначеними в повідомленні параметрами (табл. 2).

Розглянемо кілька прикладів гіпотетичних повідомлень, вихідні дані і результати розрахунків для яких представлені в таблиці 3. Для умовного повідомлення з 128 слів, згрупованих у три приклади, ми обчислимо ентропійний стан кожного повідомлення та спробуємо провести формалізовану оцінку їх валідності

Результати аналізу рівнів достовірності змодельованих повідомлень за допомогою радар-графіків (рис. 3), на яких зображені всі три версії наших прикладів, очевидні. Чим більша «площа» фігури на графіку, тим різноманітніша та збалансованіша інформація є у повідомленні. Інформативність графіків полягає в наступному.

1. Варіант 3 повідомлення має компактну, лаконічну форму – ознака як мінімум інформаційної одноманітності. Він вирізняється дуже низькою хаотичністю та відносно високою банальністю – це ознака стереотипного, банального тексту.

2. Друга версія повідомлення має найбільш рівномірний і найширший показник радар-фігури, що свідчить про високу достовірність інформації. Вона має найвищу когерентність і ентропію, а також збалансовану хаотичність, що в сукупності дає сигнал правдоподібності.

3. У тексті першого варіанту показані помірні значення за всіма показниками, що підтверджують його проміжну позицію.

На завершення розглянемо текст реального повідомлення, взятого зі стрічки новин українського контенту Obvoz.ua за 01.06.2025. У повідомленні йдеться про питання енергозабезпечення України влітку 2025 року та про те, до чого українцям варто готуватися. Обсяг тексту становив 299 слів, включаючи 7 видів груп з однаковими словами по два, три і т.д. слів. Вихідні дані та результати розрахунку тексту на неправдивість інформації представлені в таблиці 3 (остання колонка).

Таблиця 2

Чек-лист для формалізованої оцінки рівнів довіри текстового повідомлення (збільшення розміру шрифту в таблиці визначає рівень впевненості кожного наступного кластера)

Оцінка ентропії, у.о.	Параметрична оцінка повідомлення на довіру до тексту
Оцінка ентропії (Ознакою неправдивої інформації є її низька ентропія)	
0,0–0,2	Одноманітність фраз – ознака брехливості всієї інформації
0,2–0,4	Частотність вживання різних слів приблизно однакова
0,4–0,7	Різноманіття слів знаходиться на хорошому рівні, вселяє довіру до тексту
0,7–1,0	Унікальність словосполучень, різноманіття мови, як основа високої довіри до тексту
Оцінка когерентності (Чим більше ідей і пропозицій пов'язані між собою, тим вище зв'язність)	
0,0–0,2	Нелогічний текст, розбивки в сенсі і структурі – основа фейкового викладу
0,2–0,4	Логіка присутня, але є упущення або нестиковки
0,4–0,7	Ідеї логічно пов'язані між собою, текст структурований і довірчий
0,7–1,0	Дуже зв'язний і логічний текст, ідеально вибудовані зв'язки між реченнями та ідеями, вселяє впевненість у своєму значенні
Оцінка банальності (чим менш банальності, тем більш оригінальний текст)	
0,0–0,2	Недовіру викликають заїзні кліше, фрази, що часто повторюються
0,2–0,4	Повторювані фрази зустрічаються не надто часто, але вони викликають сумніви
0,4–0,7	Недовіру викликає велика кількість шаблонних виразів
0,7–1,0	Висока довіра до тексту з унікальними фразами та без шаблонів
Оцінка хаотичності (Чим менш хаотичний текст, тим він більш структурований і логічний)	
0,0–0,2	Чітка структура, речення, які плавно переходять один в одного – основа довіри до такого тексту
0,2–0,4	Є невеликі деформації в структурі, але в цілому текст логічний
0,4–0,7	Деякі фрагменти важкі для розуміння, але в цілому текст здається зрозумілим
0,7–1,0	Інформація здається неправдивою через хаос, уривчасті фрази з нелогічними переходами
Семантична новизна (Чим оригінальніший текст, тим вища його новизна)	
0,0–0,2	Текст викликає сумніви через повторюваність відомих ідей і не приносить нічого нового
0,2–0,4	Ідеї частково нові, але в основному знайомі
0,4–0,7	Пропонуються нові ідеї або аргументи, текст виходить цікавим і оригінальним, що вселяє впевненість в його об'єктивності
0,7–1,0	Дуже оригінальний і новий текст, приносить нові ідеї та концепції, що вселяє до себе впевненість в об'єктивності

Таблиця 3

Порівняльні дані інформаційно-ентропійного аналізу умовного повідомлення, що складається з 128 слів (під четвертим варіантом мається на увазі реальне текстове повідомлення)

Показник	Варіант 1	Варіант 2	Варіант 3	Oboz.ua
Структура груп слів	$12 \times 2; 5 \times 3;$ $5 \times 4; 2 \times 5$	$25 \times 2; 1 \times 3;$ $1 \times 4; 2 \times 5$	$25 \times 2; 25 \times 3$	$128 \times 1; 34 \times 2; 8 \times 3; 7 \times 4; 4 \times 5; 3 \times 6; 1 \times 13$
Число унікальних слів	59	61	3	185
Ентропія Шеннона (H)	6,14	6,33	5,66	7,228
Когерентність, $N(H)$	6,64 ($H_{opt} = 6,5; C = 7$)	6,83	6,16	0,6789
Банальність, $B(H)$	0,00216	0,00178	0,00345	7,5314
Дисперсія ймовірностей D	0,60	0,55	0,30	0,0272
Хаотичність повідомлення, $C(H)$	-3,684	-3,480	-1,698	-0,000149
Інтерпретація вірогідностей	Помірно правдоподібне, частково шаблонне	Найбільш природне, збалансоване	Маловірогідне, кліше-синтетичне	З ймовірністю 90 % комунікації є об'єктивними, а не фейковими

Результатом такого аналізу є судження про те, що цей звіт від 01.06.2025 Oboz.ua є об'єктивним і далеко не хибним, принаймні в обсязі даних, які викладені в повідомленні.

Висновки

Запропонована методика оцінки рівнів неправдивості текстових повідомлень за прямими і непрямими показниками, незважаючи на її формалізацію, дозволяє в першому наближенні оцінити достовірність конкретних повідомлень, що дозволяє більш об'єктивно підходити до інформації в мережі Інтернет, яка вже довгий час обходиться без критики або цензури, і має стійку тенденцію до розвитку фальшивості цілих інформаційних потоків. Такі підходи можуть бути корисними, зокрема, в наукових публікаціях, при рецензуванні наукових статей і книг.

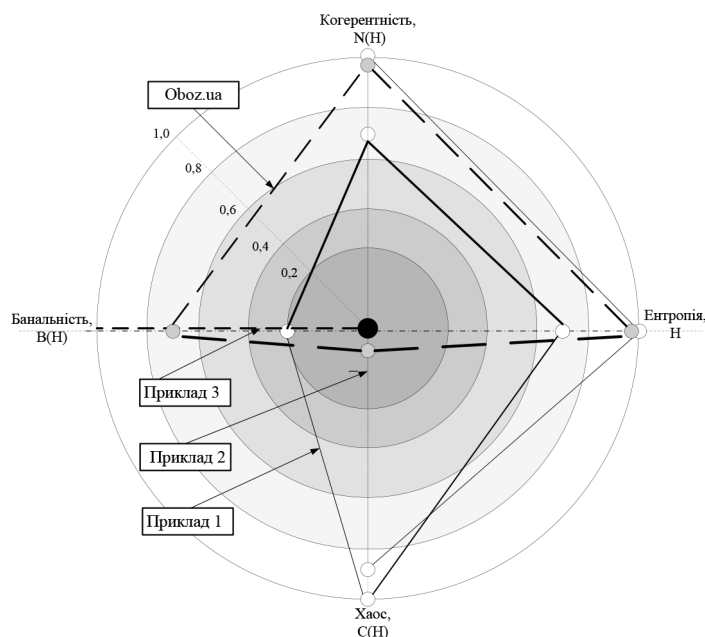


Рис. 3. Радар-графік для оцінки стану формалізованих повідомлень, що складаються з 128 слів в різних групах початкових сигналів (слів)

Список використаної літератури

1. Yavorska O. "The Knowledge As a Complementary Part of the Intellectual Capital of the Enterprises in the Era of the Digital Economy". Herald of Economics, no. 3(97), 2020, P. 185-198. – DOI: 10.35774/visnyk2020.03.185
2. Aral V.R. The spread of true and false news online. Science, 2018. DOI: 10.1126/science.aap9559
3. Reinsel D., Gantz J., Rydning J. The Digitization of the World. From Edge to Core. IDC White Paper I., (официальный идентификатор US44413318), 2018. P. 11.
4. Zhang C., Gupta A., Kauten C., Deokar A. V., Qin, X. Detecting Fake News for Reducing Misinformation Risks Using Analytics Approaches. European Journal of Operational Research (Elsevier, Amsterdam), 279(3), 2019. pp. 1036–1052. – DOI: 10.1016/j.ejor.2019.06.022
5. Shu K. Liu, H., Detecting Fake News on Social Media. Synthesis Lectures on Data Mining and Knowledge Discovery (Morgan & Claypool, San Rafael, CA), 2019, xiii + 115 pp. – DOI: 10.1007/978-3-031-01915-9
6. Shu, K., Wang S., Liu H. A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. ACM Computing Surveys (Association for Computing Machinery, New York, NY), 53(5), Article 109, 2020. P. 109:1–109:40. – DOI: 10.1145/3395046
7. Vosoughi S., Roy D., Aral S. The Spread of True and False News Online. Science (AAAS, Washington, DC), 359(6380), 2018. P. 1146–1151. – DOI: 10.1126/science.aap9559
8. Pennycook G., Rand D. G. Fighting Misinformation on Social Media Using Crowdsourced Judgments of News Source Quality. Proceedings of the National Academy of Sciences (PNAS) (National Academy of Sciences, Washington, DC), 116(7), 2019. P. 2521–2526. – DOI: 10.1073/pnas.1806781116
9. Cafaro C., Giffin A., Ali S.A., Kim D.-H. Reexamination of an Information Geometric Construction of Entropic Indicators of Complexity. 2010. 20 p. – DOI: 10.1016/J.AMC.2010.08.028
10. Hall M. J. W. Universal Geometric Approach to Uncertainty, Entropy and Information. 1999. 15 p. – DOI: 10.1103/PhysRevA.59.2602
11. Bertani A., Mazzeo V., Gallotti R. – Decoding the News Media Diet of Disinformation Spreaders. Entropy (MDPI), Vol. 26, Basel, 2024, Article 270 – DOI: 10.3390/e26030270
12. Gray R. Entropy and Information Theory., USA, Springer, 1990. 412 p. – DOI: 10.1007/978-1-4419-7970-4
13. Blasch E., Schuck T., Gagne O.B. Trusted Entropy-Based Information Maneuverability for AI Information Systems Engineering. Engineering Artificially Intelligent Systems, Lecture Notes in Computer Science (LNCS), Springer, Cham, 2021, pp. 34–52. – DOI: 10.1007/978-3-030-89385-9_3

References

1. Yavorska O. "The Knowledge As a Complementary Part of the Intellectual Capital of the Enterprises in the Era of the Digital Economy". Herald of Economics, no. 3(97), 2020, P. 185–198. – <https://doi.org/10.35774/visnyk2020.03.185>
2. Aral V.R. The spread of true and false news online. Science, 2018. <https://doi.org/10.1126/science.aap9559>

3. Reinsel D., Gantz J., Rydning J. The Digitization of the World. From Edge to Core. IDC White Paper I., (official identifier US44413318), 2018. P. 11.
4. Zhang C., Gupta A., Kauten C., Deokar A. V., Qin, X. Detecting Fake News for Reducing Misinformation Risks Using Analytics Approaches. European Journal of Operational Research (Elsevier, Amsterdam), 279(3), 2019. pp. 1036–1052. – <https://doi.org/10.1016/j.ejor.2019.06.022>
5. Shu K. Liu, H., Detecting Fake News on Social Media. Synthesis Lectures on Data Mining and Knowledge Discovery (Morgan & Claypool, San Rafael, CA), 2019, xiii + 115 pp. – <https://doi.org/10.1007/978-3-031-01915-9>
6. Shu, K., Wang S., Liu H. A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. ACM Computing Surveys (Association for Computing Machinery, New York, NY), 53(5), Article 109, 2020. P. 109:1–109:40. – <https://doi.org/10.1145/3395046>
7. Vosoughi S., Roy D., Aral S. The Spread of True and False News Online. Science (AAAS, Washington, DC), 359(6380), 2018. P. 1146–1151. – <https://doi.org/10.1126/science.aap9559>
8. Pennycook G., Rand D. G. Fighting Misinformation on Social Media Using Crowdsourced Judgments of News Source Quality. Proceedings of the National Academy of Sciences (PNAS) (National Academy of Sciences, Washington, DC), 116(7), 2019. P. 2521–2526. – <https://doi.org/10.1073/pnas.1806781116>
9. Cafaro C., Giffin A., Ali S. A., Kim D.-H. Reexamination of an Information Geometric Construction of Entropic Indicators of Complexity. 2010. 20 p. – <https://doi.org/10.1016/J.AMC.2010.08.028>
10. Hall M. J. W. Universal Geometric Approach to Uncertainty, Entropy and Information. 1999. 15 p. – <https://doi.org/10.1103/PhysRevA.59.2602>
11. Bertani A., Mazzeo V., Gallotti R. – Decoding the News Media Diet of Disinformation Spreaders. Entropy (MDPI), Vol. 26, Basel, 2024, Article 270 – <https://doi.org/10.3390/e26030270>
12. Gray R. Entropy and Information Theory., USA, Springer, 1990. 412 p. – <https://doi.org/10.1007/978-1-4419-7970-4>
13. Blasch E., Schuck T., Gagne O.B. Trusted Entropy-Based Information Maneuverability for AI Information Systems Engineering. Engineering Artificially Intelligent Systems, Lecture Notes in Computer Science (LNCS), Springer, Cham, 2021, pp. 34–52. – https://doi.org/10.1007/978-3-030-89385-9_3

Дата першого надходження рукопису до видання: 20.09.2025

Дата прийнятого до друку рукопису після рецензування: 15.10.2025

Дата публікації: 28.11.2025