

---

## INFORMATION TECHNOLOGY

---

DOI: 10.20535/2411-1031.2018.6.1.152899

УДК 004.942

АРТЕМ СОБОЛЄВ,

ДМИТРО ЛАНДЕ

### РАНЖУВАННЯ ВУЗЛІВ КВАЗІЄРАРХІЧНИХ МЕРЕЖ СОЦІАЛЬНОГО ХАРАКТЕРУ ЗА ЛОГАРИФМІЧНИМИ ВАГАМИ ЇХ ЗВ'ЯЗКІВ

Мережі соціального характеру є основою всіх взаємодій між учасниками (зазвичай, людьми), що відбуваються в процесі передавання інформації між ними. За останній час використання таких мереж все більше набирає популярності. За оцінкою експертів у даних мережах можливо отримати багато цінної інформації про їх учасників: від персональних даних (місце проживання, контактний номер телефону, адресу електронної скриньки) до поведінки та оцінки смаків. При оцінюванні мережі соціального характеру представляються графами, що відображають реальну взаємодію вузлів у таких мережах. Для оцінювання потрібно відранжувати вузли, провести їх детальний аналіз, оцінити отримані результати за всіма заданими критеріям та виділити серед них найбільш впливові. Існуючі алгоритми ранжування в переважній більшості оцінюють тільки ієрархічні мережі, що не дозволяє чітко визначити впливовість вузлів між собою. Велика кількість зв'язків між вузлами у мережах соціального характеру та їх різноманітний розподіл в більшості випадків не дозволяє використовувати алгоритми ранжування базового виду тому, що неврахування незначних, на перший погляд, зв'язків призводить до хибних результатів. У роботі приводиться аналіз процесу ранжування вузлів та порівнюються отримані результати роботи відомих алгоритмів ранжування вузлів HITS, PageRank з експертною оцінкою мережі. Для ефективності дослідження квазіієрархічних мереж, модифіковано базовий алгоритм HITS, що дозволяє проводити аналіз та ранжування вузлів за заданим критеріям (кількості вхідних і вихідних зв'язків між собою). Це узгоджується з результатами експертної оцінки. Показано, що завдяки використанню модифікованого алгоритму отримуються результати, що відповідають реальним соціальним зв'язкам між суб'єктами, а показники авторства вузлів – попередньо наданим соціальним ролям. Отриманий алгоритм дозволяє оцінити і виявляти найбільш значимі вузли в мережах соціального характеру та може використовуватися у різних галузях, де формуються соціальні мережі.

**Ключові слова:** квазіієрархічні мережі соціального характеру, суб'єкт мережі, алгоритм ранжирування, алгоритм HITS, алгоритм PageRank, ранжування вузлів, F-міра, експертна оцінка, закон Ципфа.

**Постановка проблеми.** Соціальні мережі поступово переростають в складову суспільства і являються одним із механізмів впливу на нього. Мережі соціального характеру дозволяють нав'язати думку та впливати на суб'єктів мережі через лідерів в даних мережах, тому потрібно чітко розуміти можливі ризики і наслідки у разі недбалого ставлення до даної ситуації.

Для створення механізмів контролю в мережах соціального характеру необхідно проаналізувати напрями пов'язані з протидією інформаційного впливу у мережах.

У наукових дослідженнях виокремлюють такі напрями:

- створення систем моніторингу соціальних ресурсів мережі Інтернет, з метою розвідки та інформаційного протиборства;
- розроблення методів і алгоритмів проведення інформаційних операцій у відкритих (закритих) ресурсах Інтернету.

Процес дослідження даних в соціальних мережах обумовлений швидким розвитком мережі, для дослідження якої потрібний ефективний алгоритм ранжування, який дозволить визначити найбільш критичні вузли в мережах соціального характеру.

**Аналіз останніх досліджень і публікацій.** У інформаційно-технологічному середовищі останніх 5 років активно триває процес дослідження соціальних мереж, визначається їх будова, цінність та аналізується процес впливу на них. Вивчення механізмів, які дозволяють контролювати мережі та проводити з ними певні маніпуляції дуже необхідні в умовах інформаційного нав'язування, того що, вигідно в першу чергу маніпулятору, починаючи з продажу непотрібних речей та закінчуючи інформаційними війнами. Процес аналізу соціальних мереж являється відповідальним, адже не правильна дія може призвести до неочікуваного результату або в гіршому випадку – втрати контролю над цією мережею. Проблеми ефективного дослідження і оцінювання мереж соціального характеру досліджували у своїх роботах [1], [9] та моделей інформаційного впливу в соціальних мережах [6], [8] з метою вироблення управлінських рішень щодо інформаційного протиборства в соціальних мережах.

**Метою роботи** є підвищення ефективності ранжування вузлів квазіієрархічних мереж соціального характеру шляхом модифікації алгоритму HITS і його порівняння з іншими алгоритмами ранжування вузлів в даних мережах.

**Виклад основного матеріалу дослідження.** Під ранжування мережі соціального характеру потрібно розуміти процес упорядкування вузлів (соціальних суб'єктів) за конкретною ознакою, який дозволяє визначити впливовість заданих вузлів між собою. Для візуалізації дану мережу зображають у вигляді направленої графа. В ньому вузлами відображають маркери суб'єктів (профілі соціальних мереж, телефонні номери, адреси електронних скриньок), а ребрами – направлені зв'язки між ними. Даний граф представляють матрицею суміжності вузлів. Тоді як значення у клітинках даної матриці відображає кількість направлених зв'язків.

Згідно з алгоритмом HITS, для кожного вузла мережі розраховується два показники:

- показник посередництва (*hub*);
- показник авторства (*auth*).

Показник посередництва  $hub A_i$  вузла  $A_j$  мережі, що складається із  $n$  вузлів дорівнює сумі значень авторства вузлів, на які він посилається:

$$hub A_i = \sum_{j \leftarrow i} auth A_j,$$

де  $auth A_j$  – показник авторства вузла  $A_j$  мережі;

$i$  – поточний вузол ранжування;

$j$  – вузол зв'язаний з вузлом  $i$ .

Показник авторства  $auth A_i$  вузла  $A_j$  дорівнює сумі значень посередництв вузлів, які посилаються на нього:

$$auth A_i = \sum_{j \rightarrow i} hub A_j,$$

де  $hub A_j$  – показник посередництва вузла  $A_j$  мережі.

$i$  – поточний вузол ранжування;

$j$  – вузол на який вказує вузол  $i$ .

Алгоритм HITS задає швидко збіжний ітеративний метод розрахунку показників посередництва та авторства для усіх вузлів мережі.

Водночас, коли мова йде про більш складні мережі (безмасштабні мережі), де кількість вузлів може становити більше 1000 (великі організації), складно без використання спеціальних алгоритмів робити припущення щодо соціальних суб'єктів, які виконують роль керівників різного рангу [1].

Безмасштабним мережам властива така характеристика як розподіл ступенів вузлів, який визначається як імовірність того, що вузол має ступінь (ступінь вузла – це кількість ребер, зв'язаних з цим вузлом). Саме безмасштабні (scale-free) мережі часто спостерігаються у реально існуючих мережах. Зокрема, більшість соціальних мереж є безмасштабними [2].

Щоб визначити впливовість вузлів і продемонструвати застосування базового алгоритму HITS (Hyperlink-Induced Topic Search) розглянемо звичайну ієрархічну мережу з 6 вузлами, яка представлена на рис. 1, і кожному вузлу прорахуємо коефіцієнти авторства та посередництва.

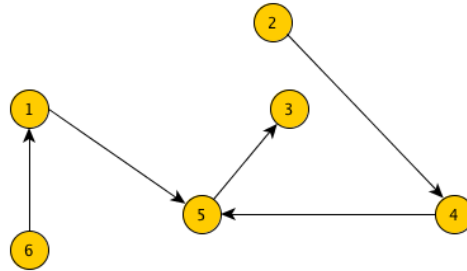


Рисунок 1 – Ієрархічна мережа

Зобразимо зв'язки між вузлами даного графу у вигляді матриці суміжності:

$$L = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Після транспонування отримуємо наступну матрицю:

$$L^T = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Для розрахунку за базовим алгоритмом HITS показників авторства та посередництва за даними, що представлені матрицями, використаємо наступні ітераційні формули [3]:

$$\begin{aligned} auth^{(k)} &= L^T hub^{(k-1)}, \\ hub^{(k)} &= L auth^{(k)}, \\ k &= k + 1, \end{aligned}$$

де  $auth^{(k)}$ ,  $hub^{(k)}$  – вектори, які містять коефіцієнти авторства і посередництва для кожної ітерації  $k$ .

Модифікуємо зазначені формули для забезпечення можливості розрахунку показників авторства та посередництва незалежно один від одного [3]:

$$\begin{aligned} auth^{(k)} &= L^T L auth^{(k-1)}, \\ hub^{(k)} &= L L^T hub^{(k)}, \\ k &= k + 1, \end{aligned}$$

Для ініціалізації даних формул як початкові матриці  $auth^{(0)}$ ,  $hub^{(0)}$  використаємо одиничні вектори:

$$L^T L = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$LL^T = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$k = 1,$

$$auth(k) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 2 \\ 0 \end{bmatrix},$$

$$hub(k) = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 2 \\ 1 \\ 1 \end{bmatrix},$$

За отриманими результатами ранжування вузлів базовим алгоритмом NITS зробимо висновок, що п'ятий вузол містить найбільш релевантну інформацію, та вузли під номером 2 і 4 являються найбільш портативні. Вони мають посилання на вузол, що містить найбільш корисну інформацію.

Розглянутий граф складається з малої кількості зв'язків і його, можливо наглядно відранжирувати. В даній мережі найбільш впливовим вузлом є вузол під номером 5, оскільки він має найбільшу кількість вхідних і вихідних зв'язків.

Можна відмітити, що кількість зв'язків у квазіієрархічних мережах соціального характеру (КМСХ) дорівнює кількості вузлів.

$$N_i = O(N_j),$$

де  $N_i$  – кількість зв'язків в мережі;

$N_j$  – кількість вузлів в мережі.

У мережах соціального характеру спостерігається три загальних варіанта зв'язків для пари вузлів:

- керівник-підлеглий;
- вузли однакового рангу;
- вузли нижчого рангу мають прямий зв'язок з вузлами більш високого рангу.

У випадку, коли вузол  $A$  є керівником вузла  $B$  в КМСХ для ранжування потрібно розглядати сумарну вагу всіх факторів, одним з яких – часовий показник.

Необхідно зазначити, що керівник може хаотично зв'язується з підлеглим, у разі необхідності, а підлеглий з певною закономірною періодичністю звітує керівникові.

При оцінці характеристики кількості зв'язків, має місце той факт, що керівник частіше зв'язується із підлеглим, тобто вага вихідного зв'язку від керівника повинна бути більшою за вагу вхідного зв'язку [1]. Потрібно зазначити, що розглядається лише узагальнений та найбільш розповсюджений шаблон контактів підлеглого та керівника. У реальній роботі цей шаблон може бути доповнений іншими специфічними характеристиками, які збільшать ефективність оцінки зв'язків між вузами [4].

Також є випадки зв'язків коли працівник нижчого рангу допомагає у робочих процесах працівникові більш високого рангу або керівник вищого рангу ставить задачу працівникові значно меншого рівня ієрархії. Це призводить до створення в такій мережі неієрархічних зв'язків, тому мережу розглядаємо як квазіієрархічну. Поміж тим, використання базового алгоритму HITS для ранжування вузлів в таких мережах не забезпечує правильного результату, оскільки навіть слабкий та незначний зв'язок (якщо не врахувати його ваги) може сильно вплинути на показник посередництва та авторства вузлів [1].

Для ранжування вузлів у мережах соціального характеру також використаємо алгоритм PageRank та порівняємо отримані результати:

$$PR(A) = (1 + d) + d \sum_{i=1}^n \frac{PR(T_i)}{C(T_i)},$$

де  $PR(A)$  – PageRank для вузла  $A$ ;

$d$  – коефіцієнт загасання;

$PR(T_i)$  – PageRank для вузла, що має посилання на вузол  $A$ ;

$C(T_i)$  – число посилань із вузла  $A$ ;

$n$  – кількість вузлів, що мають посилання на вузол  $A$  [5].

Для математичної реалізації ітеративного принципу ранжування вузлів у великих КМСХ потрібно зазначити, що найбільш впливовим у мережі суб'єктом є той, хто має найбільшу кількість вихідних зв'язків з іншими суб'єктами мережі.

Для демонстрації та визначення ефективності ранжування вузлів розглянемо КМСХ (див. рис. 2), яка складається з 14 суб'єктів між якими завідомо розподілені ролі та використаємо вищезазначені алгоритми і порівняємо отримані результати.

При ранжуванні вузлів за алгоритмом HITS отримуємо два параметри: показник авторства і показник посередництва, що не дозволяє відранжувати мережу одночасно за двома критеріями. Необхідно використати метрику, яка об'єднує у собі інформацію про показники авторства та посередництва.

Як формальну метрику для оцінки результату ранжування пропонується використати F-міра. Це дозволяє звести до одного числа два основоположних параметри: авторство і посередництво. У результаті цього отримуємо:

$$F(A_i) = \frac{2}{\frac{1}{auth A_i} + \frac{1}{hub A_i}},$$

де  $F(A_i)$  – F-міра вузла  $A_i$  мережі;

$hub A_i$  – показник посередництва;

$auth A_i$  – показник авторства.

Проаналізувавши отримані результати за табл. 1 наглядно видно, що параметри авторства і посередництва рівні між собою тому, що в КМСХ, в основному, спостерігаються зворотні зв'язки між вузами і алгоритм HITS не дозволяє правильно їх відранжувати. Алгоритм PageRank враховує тільки вхідні зв'язки і не виділяє такого параметру, як посередник, тому незначні зв'язки в мережі можуть впливати на результат ранжування. Також, порівнявши результати роботи алгоритмів HITS та PageRank з даними оцінки експертів можна стверджувати, що вищезазначені алгоритми є неефективними для ранжування вузлів в КМСХ.

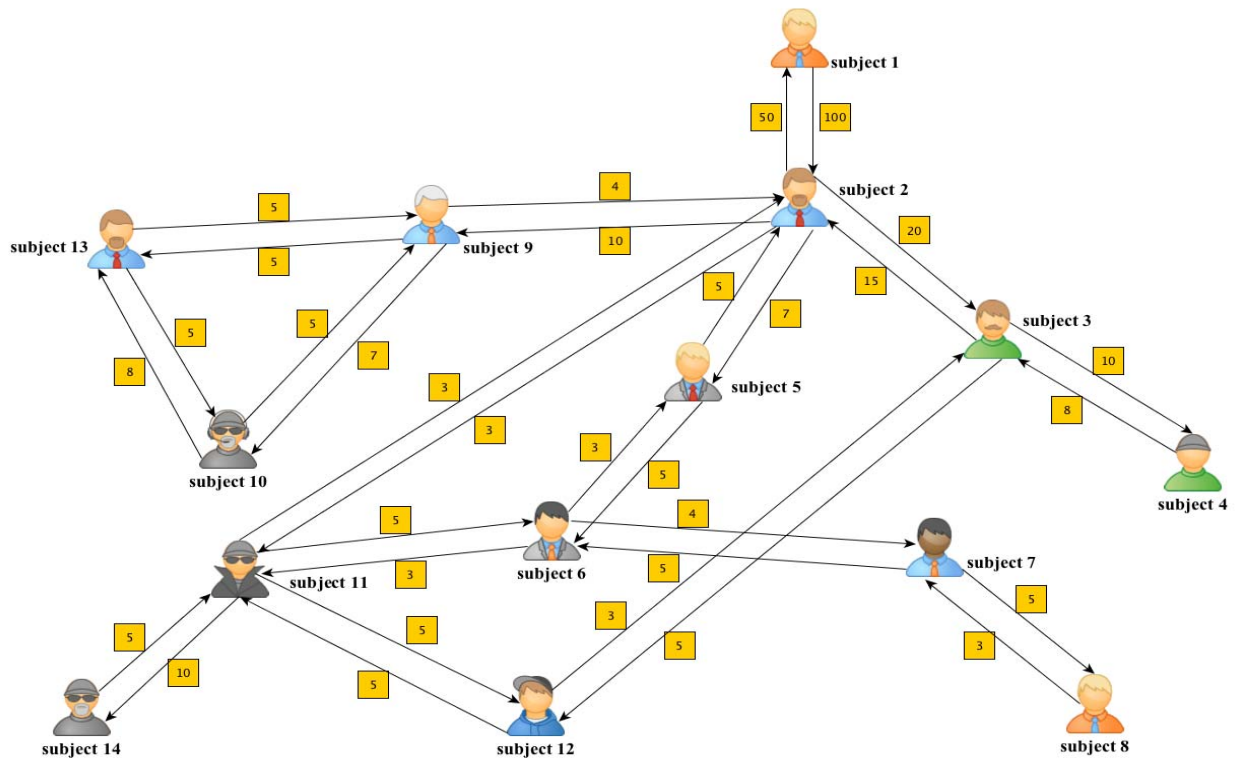


Рисунок 2 – Квазієрархічна мережа соціального характеру

Таблиця 1 – Результати ранжування вузлів за алгоритмами PageRank, HITS, та оцінкою експертів

Оцінка експертів (номер вузла)	PageRank (номер вузла)	HITS		
		auth (номер вузла)	hub (номер вузла)	F (номер вузла)
2	2	2	2	2
1	11	11	11	11
3	1	9	9	9
9	3	3	3	3
11	9	5	5	5
5	6	6	6	6
12	10	12	12	12
6	13	1	1	1
10	7	10	10	10
13	14	13	13	13
4	12	7	7	7
14	8	14	14	14
7	5	4	4	4
8	4	8	8	8

Для ефективного ранжування вузлів в КМСХ пропонується використати алгоритм HITS і при розрахунку кожного показника вузлів мережі врахувати вагу ребер. Модифікуємо алгоритм HITS шляхом помноження на вагу зв'язків між вузлами в результаті чого отримаємо:

$$hub A_i = \sum_{j \leftarrow i} auth A_j \cdot E_{ij},$$

де  $E_{ij}$  – вага зв'язків між вузлами  $A_i$  та  $A_j$ .

$$auth A_i = \sum_{j \rightarrow i} hub A_j \cdot E_{ji},$$

де  $E_{ji}$  – вага зв'язків між вузлами  $A_j$  та  $A_i$ .

Оцінивши отримані результати роботи модифікованого алгоритму HITS за табл. 2 та порівнявши з оцінкою експертів, можна стверджувати, що 7 вузлів визначено не вірно, що становить 50% від загальної кількості вузлів, тому модифікований метод HITS не є достатньо ефективним. Для детальної оцінки мережі потрібно враховувати навіть незначні зв'язки, оскільки вони можуть суттєво вплинути на результат ранжування і визначення найбільш значимих вузлів мережі.

Таблиця 2 – Результати ранжування вузлів мережі за модифікованим алгоритмом HITS та оціною експертів

Оцінка експертів (номер вузла)	HITS		
	auth (номер вузла)	hub (номер вузла)	F (номер вузла)
2	2	1	1
1	1	2	2
3	3	3	3
9	9	5	9
11	5	9	5
5	11	11	11
12	4	4	4
6	12	10	12
10	14	12	10
13	6	13	6
4	13	6	13
14	10	7	14
7	7	14	7
8	8	8	8

Так, як нерівномірність розподілу ваги може вплинути на достовірність отриманого результату ранжування, тому їх значення необхідно зменшити шляхом множення на деяку монотонно зростаючу функцію, але менш крутішу за лінійну. Розглянувши закон Ципфа, в якому основою виступає цінність зв'язків, можна стверджувати, що для будь-якого вузла мережі соціального характеру  $n$  є цінність, яка може становити від 1 до  $1/(n - 1)$ , тоді вклад даного вузла в загальну цінність мережі буде складати:

$$1 + \frac{1}{2} + \dots + \frac{1}{n} \approx \ln(n),$$

На основі вищезазначеного закону модифікуємо алгоритм HITS шляхом використання натурального логарифму ваги зв'язків між вузлами і з цього випливає наступне:

$$hub A_i = \begin{cases} \sum_{j \leftarrow i} auth A_j \cdot \ln(E_{ij}), E_{ij} > 0 \\ 0, E_{ij} = 0 \end{cases},$$

$$auth A_i = \begin{cases} \sum_{j \rightarrow i} hub A_j \cdot \ln(E_{ji}), E_{ji} > 0 \\ 0, E_{ji} = 0 \end{cases},$$

Для оцінки ефективності роботи отриманого алгоритму та перевірки відносно іншого модифікованого алгоритму HITS та використаємо вагу зв'язків в степені одна друга в першому випадку і натуральний логарифм ваги зв'язків в іншому у модифікованих формулах  $auth A_j$  та  $hub A_i$  і перевіримо результати їх роботи на КМСХ.

Оцінивши результати роботи алгоритму за табл. 3 можна стверджувати, що при використанні натурального логарифму ваги зв'язків між вузлами, модифікований алгоритм показав результати, що відповідають реальним взаємодіям між вузлами а показник авторства достатньо точно передає ієрархічну залежність суб'єктів між собою у мережі.

Даний алгоритм являється ефективним для ранжування мереж як квазіієрархічного так і ієрархічного характеру.

Також, отриманий алгоритм дозволяє проводити оцінку критичності вузлів в мережах соціального характеру, та забезпечує швидке ранжування та дослідження мереж з великою кількістю вузлів.

Необхідно відмітити, що отриманий модифікований алгоритм HITS зі значенням натурального логарифму ваги зв'язків використаний для дослідження двадцяти різних квазіієрархічних мереж, а отримані результати показників відповідали експертній оцінці.

Таблиця 3 – Результати ранжування вузлів мережі модифікованими алгоритмами HITS (степені одна друга та натуральний логарифм ваги зв'язків) та оцінкою експертів

Оцінка експертів (номер вузла)	HITS, $\frac{1}{2}$			HITS, ln		
	auth (номер вузла)	hub (номер вузла)	F (номер вузла)	auth (номер вузла)	hub (номер вузла)	F (номер вузла)
2	2	1	1	2	2	2
1	1	2	2	1	1	1
3	3	3	3	3	3	3
9	9	11	9	9	11	9
11	5	9	11	5	9	11
5	11	5	5	11	5	5
12	12	12	12	12	10	12
6	6	6	6	6	6	6
10	13	10	10	13	12	10
13	14	13	13	4	13	13
4	4	4	4	14	4	4
14	10	7	14	10	7	14
7	7	14	7	7	14	7
8	8	8	8	8	8	8

**Висновки.** Проведено аналіз ефективності відомих алгоритмів ранжування HITS, PageRank в порівнянні з експертною оцінкою квазіієрархічної мережі соціального характеру. Визначено, що алгоритм PageRank враховує тільки вхідні зв'язки і не виділяє такого параметру, як посередник, що не дозволяє ефективно ранжувати вузли в квазіієрархічних мережах. Алгоритм HITS не включає вагу ребер і для ранжування вузлів потребувало його модифікації, шляхом використання натурального логарифму ваги зв'язків між вузлами.

Для ефективності оцінки отриманих результатів роботи алгоритму HITS за двома параметрами авторства та посередництва використано метрику F-міра, у результаті чого отримано один загальний параметр для оцінки вузла.

Модифікований алгоритм HITS дозволяє ранжувати вузли мереж соціального характеру та при практичному застосуванні отримувати результати, що відповідають реальним соціальним відношенням, а показники авторства вузлів – їх ієрархічній важливості.



## СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

- [1] Д. В. Ланде, та О. О. Нечаєв, “Алгоритм ранжирування вузлів квазіієрархічних мереж соціального характеру”, *Проблеми інформатизації та управління*, т.1, № 49, с. 46-50, 2015.  
doi: 10.18372/2073-4751.1.8917.
- [2] D. Cohen, “All the World’s a Net”, *Newscientist*, vol. 174, no. 2338, pp. 24-29, 2002.
- [3] A. N. Langville, and C. D. Meyer, *Google’s PageRank and Beyond: The Science of Search Engine Rankings*. Princeton, New Jersey, USA: Princeton university press, 2013.
- [4] Y. Y. Liu, J. J. Jean-Jacques Slotine, and A. L. Barabasi, “Control centrality and hierarchical structure in complex networks”, *PLOS ONE*, vol. 7, № 9, pp. 1-7, 2012.  
doi: 10.1371/journal.pone.0044459.
- [5] Д. В. Ландэ, А. А. Снарский, та И. В. Безсуднов, *Интернетика: навигация в сложных сетях: модели и алгоритмы*. Москва, Российская Федерация: Либроком, 2009.
- [6] О. К. Сулема, та Д. В. Ланде, “Знаходження оптимальної ієрархії у квазіієрархічному графі за критеріями центральності”, *Реєстрація, зберігання і обробка даних*, т. 17, № 4, с. 3-10, 2015.
- [7] L. Laura, and Me. Gianluigi, “Searching the Web for illegal content: the anatomy of a semantic search engine”, *Soft Computing*, vol. 21, iss. 5, pp. 1245-1252, 2017.  
doi: 10.1007/s00500-015-1857-4.
- [8] Д. А. Губанов, и А. Г. Чхартишвили, “Формальные и неформальные связи пользователей социальной сети Facebook”, на *XII Всероссийского совещании по проблемам управления*, Москва, 2014, с. 6301-6309.
- [9] А. Г. Додонов, Д. В. Ландэ, и В. Г. Путятин, *Компьютерные сети и аналитические исследования*. Киев, Украина: ИПРИ НАН Украины, 2014.
- [10] J. Scott, *Social network analysis*. London, UK: SAGE Publications Ltd, 2017.
- [11] M. R. Parks, “Personal relationship and personal networks”, *Journal of Communication*, vol. 58, iss.2, pp. 397-399, 2008.  
doi: 10.1111/j.1460-2466.2008.00391\_2.x.
- [12] S. P. Borgatti, M. G. Everett, and J. C. Johnson, *Analyzing Social Networks*. London, UK: SAGE Publications Ltd, 2018. –pp. 13-25.

Стаття надійшла до редакції 10 березня 2018 року.

## REFERENCE

- [1] D. V. Lande, and A. O. Nechaev, “Algorithm for ranking of nodes of quasi-hierarchical networks of social character”, *Problems of informatization and management*, vol. 1, no. 49, pp. 46-50, 2015.  
doi: 10.18372/2073-4751.1.8917.
- [2] D. Cohen, “All the World’s a Net”, *Newscientist*, vol. 174, no. 2338, pp. 24-29, 2002.
- [3] A. N. Langville, and C. D. Meyer, *Google’s PageRank and Beyond: The Science of Search Engine Rankings*. Princeton, New Jersey, USA: Princeton university press, 2013.
- [4] Y. Y. Liu, J. J. Jean-Jacques Slotine, and A. L. Barabasi, “Control centrality and hierarchical structure in complex networks”, *PLOS ONE*, vol. 7, no. 9, pp. 1-7, 2012.  
doi: 10.1371/journal.pone.0044459.
- [5] D. V. Lande, A. A. Snarsky, and I. V. Bezsudnov, *Internet: Navigation in complex networks: models and algorithm*. Moskow, Russia: Librokom, 2009.
- [6] O. K. Sulema, and D. V. Lande, “Finding the optimal hierarchy in the quasi-hierarchical graph according to the criteria of centrality”, *Registration, storage and processing of data*, vol. 17, no. 4, pp. 3-10, 2015.

- [7] L. Laura, and Me. Gianluigi, “Searching the Web for illegal content: the anatomy of a semantic search engine”, *Soft Computing*, vol. 21, iss. 5, pp. 1245-1252, 2017.  
doi: 10.1007/s00500-015-1857-4.
- [8] D. A. Gubanov, and A. G. Chkhartishvili, “Formal and informal communications of users of the social network Facebook”, in *XII All-Russia Meeting on the Problems of Management*, Moscow, 2014, pp. 6301-6309.
- [9] A. G. Dodonov, D. V. Lande, and V. G. Putiatin, *Computer networks and analytical studies*. Kyiv, Ukraine: IPRI NAS of Ukraine, 2014.
- [10] J. Scott, *Social network analysis*. London, UK: SAGE Publications Ltd, 2017.
- [11] M. R. Parks, “Personal relationship and personal networks”, *Journal of Communication*, vol. 58, iss.2, pp. 397-399, 2008.  
doi: 10.1111/j.1460-2466.2008.00391\_2.x.
- [12] S. P. Borgatti, M. G. Everett, and J. C. Johnson, *Analyzing Social Networks*. London, UK: SAGE Publications Ltd, 2018. –pp. 13-25.

АРТЕМ СОБОЛЕВ  
ДМИТРИЙ ЛАНДЕ

### **РАНЖИРОВКА УЗЛОВ КВАЗИИЕРАРХИЧЕСКИХ СЕТЕЙ СОЦИАЛЬНОГО ХАРАКТЕРА ЗА ЛОГАРИФМИЧЕСКИМИ ВЕСАМИ ИХ СВЯЗЕЙ**

Сети социального характера является основой всех взаимодействий между участниками (как правило, людьми), происходящие в процессе передачи информации между ними. За последнее время использование таких сетей все больше набирает популярность. По оценке экспертов в данных сетях можно получить много ценной информации об их участниках: от персональных данных (место жительства, контактный номер телефона, адрес электронной почты) к поведению и оценки вкусов. При оценке сети социального характера представляются графами, отражающие реальное взаимодействие узлов в таких сетях. Для оценки нужно проранжировать узлы, провести их детальный анализ, оценить полученные результаты по всем заданным критериям и выделить среди них наиболее влиятельные. Существующие алгоритмы ранжирования в большинстве оценивают только иерархические сети, что не позволяет четко определить влиятельность узлов между собой. Большое количество связей между узлами в сетях социального характера и их разнообразное распределение в большинстве случаев не позволяет использовать алгоритмы ранжирования базового вида потому, что неучет незначительных, на первый взгляд, связей приводит к ошибочным результатам. В работе приводится анализ процесса ранжирования узлов и сравниваются полученные результаты работы известных алгоритмов ранжирования узлов HITS, PageRank и экспертной оценке сети. Для эффективности исследования квазиерархических сетей, модифицировано базовый алгоритм HITS, что позволяет проводить анализ и ранжирование узлов по заданным критериям (количества входных и выходных связей между собой). Это согласуется с результатами экспертной оценки. Показано, что благодаря использованию модифицированного алгоритма получаются результаты, соответствующие реальным социальным связям между субъектами, а показатели авторства узлов - предварительно предоставленным социальным ролям. Полученный алгоритм позволяет оценить и выявлять наиболее значимые узлы в сетях социального характера и может использоваться в различных областях, где формируются социальные сети.

**Ключевые слова:** квазиерархические сети социального характера, субъект сети, алгоритм ранжирования, алгоритм HITS, алгоритм PageRank, ранжирование узлов, F-мера, экспертная оценка, закон Ципфа.

DMYTRO LANDE,  
ARTEM SOBOLIEV

## QUASI-HIERARCHICAL SOCIAL NETWORK NODES RANKING IN TERMS OF LOGARITHMIC FUNCTION OF ITS LINK WEIGHTS

Networks of social nature are the basis of all interactions between participants (usually people) that occur in the process of transferring information between them. Recently, the use of such networks is gaining in popularity. According to expert estimates in these networks, it is possible to get a lot of valuable information about their participants: from personal data (place of residence, contact number, e-mail address) to behavior and assessment of flavors. When evaluating a social networks, graphs representing the real interaction of nodes in such networks are represented. For evaluation, it is necessary to rotate the nodes, to conduct their detailed analysis, to evaluate the results obtained according to all the given criteria and to allocate among them the most influential. Existing algorithms for ranking in the overwhelming majority are estimated only by hierarchical networks, which does not allow to clearly determine the influence of nodes among themselves. The large number of links between nodes in social networks and their diverse distribution in most cases does not allow for the use of algorithms ranking the basic form, because the ignoring of minor, at first glance, linkages leads to false results. The paper analyzes the ranking of nodes and compares the results of the known algorithms for ranking the nodes HITS, PageRank with expert evaluation of the network. For the efficiency of the quasi-hierarchical networks research, the basic HITS algorithm was modified, which allows to analyze and rank nodes according to the given criteria (the number of input and output links among themselves). This is consistent with the results of expert evaluation. It is shown that due to the use of the modified algorithm results are obtained that correspond to real social connections between the subjects, and the indexes of the authorship of the nodes - to the previously provided social roles. The resulting algorithm allows you to evaluate and identify the most significant nodes in social networks and can be used in various industries where social networks are formed.

**Keywords:** quasi-hierarchical networks of social nature, network subject, ranking algorithm, HITS algorithm, PageRank algorithm, nodes ranking, F-measure, expert evaluation, Zipf law.

**Артем Миколайович Соболев**, аспірант, Інститут спеціального зв'язку та захисту інформації Національного технічного університету України “Київський політехнічний інститут імені Ігоря Сікорського”, Київ, Україна.

ORCID: 0000-0003-4027-042X.

E-mail: artem1sobolev@gmail.com.

**Дмитро Володимирович Ланде**, доктор технічних наук, старший науковий співробітник, завідувач відділом спеціалізованих засобів моделювання, Інститут проблем реєстрації інформації Національної академії наук України, Київ, Україна.

ORCID: 0000-0003-3945-1178.

E-mail: dwlande@gmail.com.

**Артем Николаевич Соболев**, аспирант, Институт специальной связи и защиты информации Национального технического университета Украины “Киевский политехнический институт имени Игоря Сикорского”, Киев, Украина.

**Дмитрий Владимирович Ландэ**, доктор технических наук, старший научный сотрудник, заведующий отделом специализированных средств моделирования, Институт проблем регистрации информации Национальной академии наук Украины, Киев, Украина.

**Artem Soboliev**, postgraduate student, Institute of special communication and information protection of National technical university of Ukraine “Igor Sikorsky Kyiv polytechnic institute”, Kyiv, Ukraine.

**Dmytro Lande**, doctor of technical science, senior researcher, head of the specialized modeling tools department, Institute for information recording of National academy of science of Ukraine, Kyiv, Ukraine.