

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/279535422>

ОГЛЯД МОДЕЛЕЙ АНАЛІЗУ СОЦІАЛЬНИХ МЕРЕЖ

Article · January 2015

CITATIONS

0

READS

614

2 authors, including:



[Serhiy Shtovba](#)

Vinnitsia National Technical University

104 PUBLICATIONS 240 CITATIONS

SEE PROFILE

ОГЛЯД МОДЕЛЕЙ АНАЛІЗУ СОЦІАЛЬНИХ МЕРЕЖ

¹Вінницький національний технічний університет

Розглянуто історію розвитку аналізування соціальних мереж. Детально розглянуті основні типи соціальних мереж та методи їх аналізу. Розкрито актуальність розвитку та застосування методології аналізу соціальних мереж, зокрема в сучасних інформаційних технологіях.

Ключові слова: соціальні мережі, аналіз соціальних мереж, мережі гіперпосилань, кількісні показники соціальних мереж, блокові моделі, випадкові графи.

Вступ

Соціальна мережа — це будь яка соціальна взаємодія, що може бути представлена множиною соціальних одиниць та відношенням між ними. В останнє десятиліття у житті суспільства все більшу роль відіграють онлайніві соціальні мережі, в яких соціальна взаємодія здійснюється через веб-технології. Онлайнівими соціальними мережами охоплено понад 70 % користувачів Інтернету. В Європі та США один користувач входить в середньому у 2 соціальні мережі. Деякі люди звертаються до соціальної мережі лише епізодично, а деякі – практично «живуть» в них, що інколи межує із патологією. Для когось соціальна мережа — це відпочинок, але дехто не приймає рішень без обговорення в мережі.

Достовірно встановлено, що онлайніві соціальні мережі сьогодні реально впливають на поведінку їх учасників поза мережею. Віднедавна онлайніві соціальні мережі перетворилися в елемент прихованого управління соціально-економічними системами. Масштаб такого управління варіюється від просування товарів та брендів за технологією вірусного маркетингу до організації акцій масового суспільного протесту в Тунісі, Єгипті, Великобританії, Україні, Гонконзі тощо. Переваги такого управління полягають в оперативності досягнення мети, невеликих фінансових витратах, в можливості дистанційного маніпулювання свідомістю користувачів в автоматизованому режимі та майже повним протоколюванням результатів діяльності. Мати у своїх руках такі важелі управління зацікавлені як приватні корпорації, так і політичні сили та органи державної влади.

Щоб успішно управляти соціумом, спочатку треба його знати. Тому актуальною є задача аналізу соціальних мереж з метою ідентифікації багатофакторних залежностей для прогнозування поширення і сприйняття нової інформації учасниками мережі, з подальшим виявленням біфуркаційних процесів як всередині мережі, так і зовні. Розв'язання цієї задачі створить передумови для формалізованого управління поведінкою соціально-економічних систем шляхом впливу на характеристики окремих ключових елементів соціальних мереж.

Метою статті є огляд моделей аналізу соціальних мереж, з урахуванням того, що соціальні одиниці не діють незалежно, а навпаки, впливають одна на одну. Розглядаються моделі на основі теорії графів, структурної еквівалентності та випадкових графів, виявляються перспективні шляхи їх покращення. В статті робиться акцент на аналізі онлайнівих соціальних мереж.

Базові поняття

Актор. Актор це одна персона, корпоративна або колективна соціальна одиниця. Прикладами акторів можуть бути люди у групі, департаменти в корпорації, державні установи у місті, держави у світі. Якщо всі актори одного типу, наприклад, люди у групі, тоді вони утворюють однодомну мережу. Дводомною називається мережа з двома типами акторів, наприклад, учні та вчителі у школі.

Зв'язок. Актори пов'язані один з одним соціальними зв'язками, типів яких може бути велика кількість. Кожен зв'язок визначає відношення певного типу між двома акторами, наприклад:

- оцінка однієї людини іншою (прояв дружби, бажання спілкування тощо);
- передача матеріальних ресурсів (бізнес-транзакції, кредитування чи запозичення тощо);
- спільна належність до певного явища, об'єкта чи події (участь у соціальній події, належність

до одного соціального об'єднання);

- поведінкова взаємодія (спільні розмови, листування);
- переміщення між місцями чи статусами (міграція, соціальна мобільність);
- формальні відношення, наприклад, субординація;
- біологічні відношення, наприклад, батьки — діти.

Діада (двійка). В найпростішому випадку відношення є бінарним. Воно описує зв'язок між парою акторів. Сукупність двох акторів та зв'язку між ними отримала назву діада. Аналіз діад направлений на вивчення властивостей бінарних відношень, як то визначення його симетричності чи закономірностей встановлення певних типів зв'язків.

Тріада (трийка). Підмножина з трьох акторів та всі можливі зв'язки між ними утворюють ще одну одиницю аналізу — тріаду. На основі тріад розглядаються поняття транзитивності і транзитивного балансу. Також на основі транзитивного балансу здійснюється бінарна кластеризація мережі.

Підгрупа. Діади та тріади утворюють найпростіші підгрупи. Підгрупа — це будь яка підмножина акторів та всі зв'язки між ними.

Група. Аналіз соціальних мереж може розглядати на тільки діади, тріади та підгрупи, а й цілі системи акторів. Група складається зі всіх акторів та зв'язків між ними, які були отримані з концептуальних, теоретичних чи емпіричних міркувань, і розглядається як скінченна множина соціальних одиниць, які підлягають аналізу.

Соціальна мережа. Соціальна мережа складається зі скінченної множини акторів і відношень, визначених ними.

Структурні та композиційні змінні. Структурні змінні описують зв'язки певного типу між парами акторів. Наприклад, структурною змінною може бути міра бізнес-транзакцій між корпораціями, дружби між людьми, торгівлі між країнами. Композиційні змінні описують атрибути акторів, наприклад, стать, вік, етнічне походження, географічне розташування актора тощо.

Мережі належності. Мережа що містить акторів одного типу називаються однодольними. Інший тип мереж виникає під час збору структурних даних з двох (або навіть більше) наборів сутностей. Наприклад розгляд двох наборів акторів, один з яких складається з корпорацій, а інший — з неприбуткових організацій. Можна розглянути потік фінансової допомоги з корпорацій до неприбуткових організацій. Така мережа є дводольною; вона містить два набори акторів, один з яких є донором, а інший — реципієнтом. Окремим типом дводольних мереж є мережа належності, яку утворюють множина акторів та множина подій (клуби, організації, громадські організації), до якої може відноситись актор або підгрупа акторів.

Соціальна позиція — це мітка акторів, що мають еквівалентні зв'язки з іншими членами групи.

Соціальна роль — це система зв'язків між акторами чи соціальними позиціями.

Стислий історичний екскурс

Перший напрямок аналізу соціальних мереж полягав у пошуку форм їх візуалізації. Однією із перших форм представлення соціальних мереж були соціограми (Я. Морено 1953 р.) [1—3]. Я. Морено першим поставив задачу аналізу міжособистісних відносин в малих групах. Аналізуючи структуру та динаміку малих груп, Я. Морено представив людей та зв'язки між ними у двовимірному просторі за допомогою точок та відрізків, що їх з'єднували. Точки відповідали людям, а відрізки — зв'язкам. Таке представлення, що отримало назву соціограми, часто застосовувалось для представлення структури малих груп [1], зокрема: спільноти еліт (Е. О. Laumann та F. U. Pappi, 1976 р., Е. О. Laumann та D. Knoke, 1987 р.); внутрішньо-корпоративних зв'язків (J. H. Levine, 1972 р.); структури ролей в групах (R. L. Breiger, S. A. Boorman та P. Arabie, 1975 р., R. S. Burt 1976 р., 1982 р.); шаблонів взаємодії в малих групах (A. K. Romney та K. Faust, 1982 р., L. C. Freeman, S. C. Freeman та A. G. Michaelson, 1989 р.).

Представлення соціограм в матричній формі дозволило вивчати соціальні системи математичними методами. Для опису соціальних груп почали застосовувати чисельні характеристики такі, як центральність актора та централізація групи, обопільність та взаємність, структурний баланс, транзитивність. Пізніше сформулювали математичні визначення понять соціальна група, ізолюваність, популярність, престиж, баланс, транзитивність, угруповання, підгрупа, соціальна згуртованість, соціальна позиція, соціальна роль, обопільність, взаємність, обмін, вплив, домінантність, конформізм тощо [1, 3—4].

Другий напрямок аналізу соціальних мереж отримав назву структурний баланс (Heider F. 1946, T. M. Newcomb 1953) [1, 5]. Він стосується визначення позитивних та негативних зв'язків в групі,

перевірці тезису про те, що ворог мого ворогу — мій друг. Структурний баланс сприяв вивченню транзитивності мереж з урахуванням триплетів — тернарних відношень акторів. В структурному балансі введено означення структурної еквівалентності, яку приписують акторам з однаковими множинами зв'язків в мережі (F. Lorrain та H. C. White 1971) [6]. На базі структурної еквівалентності сформовано поняття роль, соціальний статус, соціальна позиція та розроблено основи класифікації соціальних мереж [6—8].

Третій напрямок аналізу мереж базується на застосуванні математичної статистики. J. A. Devis, P. W. Holland та S. Leinhardt в серії своїх праць в 1970-х роках [1] запропонували низку ймовірнісних розподілів випадкових направлених графів для перевірки різних структурних тенденцій соціальних мереж. В 1980-х роках розроблено моделі для роботи з різними типами соціальних мереж, які дозволяють аналізувати бінарні, тернарні та багатоарні відношення. [3, 9].

Графові моделі

По-перше, теорія графів надала словник понять для математичного опису властивостей соціальних структур. По-друге, теорія графів надає математичні інструменти для кількісного оцінювання характеристик соціальних мереж. По-третє, словник понять, математичні методи та теорія графів надають змогу доводити теореми про властивості соціальних структур.

В загальному випадку вершини графа відповідають акторам соціальної мережі, а ребра графа — комунікаціям між акторами (рис. 1). Розглянемо граф G деякої соціальної мережі, який містить вершини $N = \{n_1, n_2, \dots, n_g\}$ та ребра $E = \{l_1, l_2, \dots, l_L\}$. Для комп'ютерного аналізу граф соціальної мережі представляють матриця суміжності, яку називають *соціоматрицею*. Як приклад показано соціоматрицю для графа з рис. 1:

\	n_1	n_2	n_3	n_4	n_5	n_6
n_1	0	1	1	1	0	0
n_2	1	0	1	0	0	0
$X = n_3$	1	1	0	1	0	0
n_4	1	0	1	0	0	0
n_5	0	0	0	0	0	1
n_6	0	0	0	0	1	0

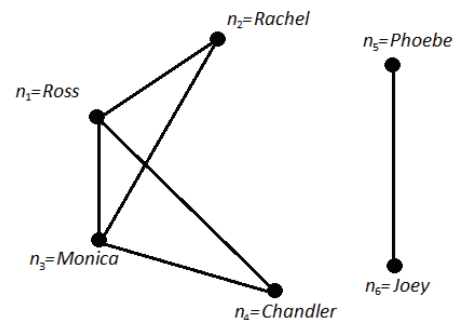


Рис. 1. Граф соціальної мережі

В аналізі соціальних мереж використовуються такі характеристики графів.

Ступінь вершини ненаправленого графа, який дорівнює кількості ребер, що їй інциденті. Для направленого графа ступінь вершини розраховують окремо для вхідних та вихідних дуг. Ступінь вершини характеризує активність актора. Узагальненим показником для мережі є *середній ступінь вершини*:

$$\bar{d} = \frac{\sum_{i=1}^g d(n_i)}{g} = \frac{2L}{g}; \quad (1)$$

$$\bar{d}_1 = \bar{d}_0 = L/g, \quad (2)$$

де $d(n_i)$ — ступінь вершини n_i ; L — кількість ребер/дуг в мережі; g — кількість вершин в мережі; \bar{d}_1 та \bar{d}_0 — середній ступінь вершини для вхідних та вихідних дуг відповідно

Середній ступінь вершини використовується в багатьох стохастичних моделях [1] як параметр активності мережі до створення зв'язків. Окрім середнього ступеню вершини використовується і оцінка його дисперсії:

$$S_D^2 = \frac{\sum_{i=1}^g (d(n_i) - \bar{d})^2}{g}, \quad (3)$$

де $d(n_i)$ — ступінь вершини n_i ; \bar{d} — середній ступінь вершин мережі; g — кількість вершин у мережі.

Щільність графа — це характеристика зв'язаності графа, яка розраховується як відношення фактичної кількості ребер до їх максимально можливої кількості:

$$\Delta = \frac{L}{g(g-1)/2} = \frac{2L}{g(g-1)}; \quad (4)$$

$$\Delta = \frac{\bar{d}}{(g-1)}. \quad (5)$$

Центральність — характеристика активності актора, яка розраховується як відношення ступеня відповідної вершини графа до кількості вершин, з якими можна утворити зв'язок:

$$C_D(n_i) = \frac{d(n_i)}{g-1}. \quad (6)$$

Централізованість групи визначається так:

$$s_C^2 = \frac{\sum_{i=1}^g (C_D(n_i) - \bar{C}_D)^2}{g}, \quad (7)$$

де $C_D(n_i)$ — дисперсія центральності акторів.

Для регулярного графа централізованість буде рівна 0, та досягати максимуму при з'єднанні типу «зірка».

Центральність актора може визначатись і у інший спосіб, наприклад, L. C. Freeman запропонував узагальнений індекс центральності та такі похідні показники як центральна близькість, центральність посередника та інформаційна центральність [1, 4].

Престиж актора визначають відношенням кількості вхідних дуг відповідної вершини графа до кількості акторів в мережі:

$$P_D(n_i) = \frac{d_1(n_i)}{g-1}. \quad (8)$$

Інколи престиж актора визначають за принципами PageRank-алгоритму з пошукової системи Google.

Однією з основних цілей застосування теорії графів є кластеризація соціальної мережі, тобто формалізоване виявлення в ній груп. Група — це фрагмент графа, щільність внутрішніх зв'язків в якому домінує над щільністю зовнішніх зв'язків. Малозв'язні групи, що мають розріджені зв'язки всередині, але водночас об'єднують декілька підгруп — є дуже критичними. Інформація в щільних групах, зазвичай, є схожою та надлишковою, тому для групи вона має меншу цінність, ніж та що проходить через слабкі зв'язки з інших підгруп [5, 10, 11].

L. C. Freeman припустив, що ідеальною підгрупою є повнозв'язна компонента графа (сильний альянс). Відповідно до таких припущень розроблено цілий ряд моделей підгруп [12]:

n-clique — максимальний підграф, в якому відстань між двома будь-якими парами вершин не більше n ;

k-plex — максимальний підграф, який містить g_s вершин та має не менше як $(g_s - k)$ суміжних вершин у підграфі;

k-core — максимальний підграф, кожна вершина якого має не менше k суміжних вершин всередині підграфа;

LS-множина — підграф, в якому довільна множина вершин має більше внутрішніх ребер, ніж зовнішніх;

λ -множина — підграф, в якому міра зв'язності будь-якої пари вершин всередині підгрупи більша, ніж поза нею. Зв'язаність вершин $\lambda(i, j)$ визначається через мінімальну кількість зв'язків, які треба видалити, щоб вершини (i, j) були недосяжні.

Пошук підгруп досить складна задача, наприклад, складність пошуку LS-множин та λ -множин дорівнює $O(n^5)$. [12]. Тому точні алгоритми виявлення підгруп можна застосувати лише для малих соціальних мереж.

Алгебраїчні моделі

Алгебраїчні моделі переважно застосовують для задач кластеризації акторів. Одна із таких задач виявлення акторів з еквівалентними зв'язками з іншими членами групи та об'єднання їх у позиції [8]. Ключові результати в цьому напрямку отримав Н. С. White [7], який формалізував поняття структурної еквівалентності. Для вимірювання структурної еквівалентності R. S. Burt [13] запропонував Евклідову відстань:

$$d_{ij} = \sqrt{\sum_{k=1}^g (x_{ik} - x_{jk})^2 + (x_{ki} - x_{kj})^2}, \quad i \neq k, \quad j \neq k, \quad (9)$$

де x_{ik} — суміжність вершин (i, k) ; g — кількість вершин в графі.

Іншою популярною мірою структурної еквівалентності є коефіцієнт кореляції:

$$r_{ij} = \frac{\sum (x_{ki} - \bar{x}_{row i})(x_{kj} - \bar{x}_{row j}) + \sum (x_{ik} - \bar{x}_{col i})(x_{jk} - \bar{x}_{col j})}{\sqrt{\sum (x_{ki} - \bar{x}_{row i})^2 + \sum (x_{ik} - \bar{x}_{col i})^2} \sqrt{\sum (x_{kj} - \bar{x}_{row j})^2 + \sum (x_{jk} - \bar{x}_{col j})^2}}, \quad (10)$$

$$i \neq k, \quad j \neq k$$

де x_{ik} — суміжність вершин (i, k) ; $\bar{x}_{row i}$ та $\bar{x}_{col i}$ — відповідно, середнє значення i -го рядка та стовпчика в соціоматриці; g — кількість вершин у графі.

Обидві міри є однаковими для абсолютно структурно еквівалентних акторів. Для інших випадків їх значення можуть сильно відрізнятись. Для визначення соціальних позицій запропоновано кілька методів кластеризації на базі мір (9)—(10), першими з яких були CONCOR та ієрархічна кластеризація.

Метод CONCOR (CONvergence of iterated CORrelation.) [7] полягає в ітераційному перерахунку кореляційної матриці з початкової матриці C_1 , елементи r_{ij} якої обчислені за формулою (10). Процедура повторюється до моменту, коли всі елементи не перетворяться на +1 або -1. Далі, перестановкою рядків та стовпчиків фінальну матрицю представляють у спрощеному вигляді блоків:

$$C_i = \begin{pmatrix} +1 & -1 \\ -1 & +1 \end{pmatrix}, \quad (11)$$

де актори кожного з блоків +1 чи -1 мають однакову соціальну позицію.

Метод ієрархічної кластеризації представляє собою жадібний алгоритм розбиття даних за деяким критерієм. На кожній ітерації алгоритму утворюється 2 підгрупи. На відміну від CONCOR, кількість отриманих внаслідок розбиття соціальних позицій залежить від умов зупинки ітераційного алгоритму та обраного критерію схожості акторів. Схожість акторів можна визначити як за кореляційною матрицею C_1 , так і за матрицею евклідових відстаней.

Розглянемо граф G з рис. 1 з трьома еквівалентними групами акторів. За виразом (9) розрахуємо відстань між вершинами графа:

\	n_1	n_2	n_3	n_4	n_5	n_6
n_1	-	$\sqrt{2}$	0	$\sqrt{2}$	$\sqrt{8}$	$\sqrt{8}$
n_2	$\sqrt{2}$	-	$\sqrt{2}$	0	$\sqrt{6}$	$\sqrt{6}$
$d = n_3$	0	$\sqrt{2}$	-	$\sqrt{2}$	$\sqrt{8}$	$\sqrt{8}$
n_4	$\sqrt{2}$	0	$\sqrt{2}$	-	$\sqrt{6}$	$\sqrt{6}$
n_5	$\sqrt{8}$	$\sqrt{6}$	$\sqrt{8}$	$\sqrt{6}$	-	0
n_6	$\sqrt{8}$	$\sqrt{6}$	$\sqrt{8}$	$\sqrt{6}$	0	-

З матриці відстаней видно, що відстань між структурно еквівалентними акторами рівна нулю. Структурно еквівалентні актори утворюють підгрупи: $\beta_1 = \{Ross, Monica\}$; $\beta_2 = \{Chandler, Rachel\}$; $\beta_3 = \{Joey, Phoebe\}$.

Зв'язність груп описується щільністю, яка характеризує ступінь зовнішніх зв'язків членів груп.

Щільність між групами β_i та β_j розраховується так:

$$\Delta_{ij} = \frac{\sum x_{kl}}{2g_i g_j}, \quad k, l \in \beta_i, \beta_j. \quad (12)$$

Інформацію про щільність мережі зручно задавати матрицею щільності $D = [\Delta_{ij}]$, $i = 1 \dots m$, $j = 1 \dots m$. Матриця щільності для груп з попереднього прикладу є такою:

$$D = \begin{array}{c|ccc} & \beta_1 & \beta_2 & \beta_3 \\ \hline \beta_1 & 1 & 1 & 0 \\ \beta_2 & 1 & 0 & 0 \\ \beta_3 & 0 & 0 & 1 \end{array}. \quad (13)$$

З (13) видно, що групи β_1 та β_2 мають високу щільність зв'язків між акторами, отже можна говорити про комунікацію між групами, на відміну від ізольованої групи β_3 . Щільність набуває значень в межах $[0; 1]$, тому однозначно ідентифікувати комунікацію між групами важко. Для визначення пов'язаних груп на основі матриці щільності з неперервними значеннями використовують блокові моделі (Н. С. White, S. A. Boorman і R. L. Breiger 1976 [7]). За методом блокових моделей матрицю щільності D перетворюють в уявну матрицю M з бінарних елементів, які слугують індикаторами зв'язку між групами. Бінаризація матриці D здійснюється за деяким пороговим значенням зв'язків всієї групи або середньої щільності зв'язків між групами.

Блокові моделі дозволяють формалізувати соціальної ролі як узагальнений шаблон відношень між позиціями. Основним об'єктом дослідження соціальної ролі є композиція відношень. Наприклад, соціальна мережа містить відношення A — допомога та F — дружба. Тоді шляхом композиції відношень, можна дослідити нові типи відношення, наприклад, $F \circ F$ — «друг мого друга» або $A \circ A$ — «допомога тому, хто уже допомагає комусь». Таким чином, соціальну роль можна визначити деякою композицією відношень, що має змістовну інтерпретацію. Відповідно, аналіз соціальної мережі спирається на досліджені композиції відношення, знаходження шаблонів — типових структур зв'язків та їх формалізацію. Деталі описаного підходу викладено в [1, 2, 8].

Стохастичні моделі

Стохастичні моделі переважно застосовують для великих соціальних мереж з тисяч і навіть мільйонів акторів. Аналіз таких мереж полягає не у знаходженні одного ключового актора, а у виявленні долі ключових акторів в мережі [14]. Цікавою є задача оцінювання відмовостійкості мережі шляхом визначення долі акторів, вилучення яких розіб'є її на кілька фрагментів чи суттєво збільшить середню відстань між акторами [15].

Стохастичні моделі поділяють на 2 групи: моделі статички та моделі динаміки. Моделі статички розглядають зв'язки як імовірнісні події з урахуванням неповноти інформації про структуру мережі та вплив на неї суб'єктивних факторів [2]. Динамічні моделі розглядають мережу з точки зору багатоагентних систем, в якій соціальні актори (активні адаптивні агенти) здатні приймати рішення, що приводять до зміни структури мережі. Соціальні й пізнавальні процеси, які можуть впливати на зв'язки в мережі, детально розглядаються в [5, 16].

Стохастичні моделі статички соціальних мереж базуються на теорії випадкових графів. Випадковим називається граф, в якому число вершин та ребер, а також їх інцидентність визначаються деяким випадковим розподілом. В дослідженнях [1, 3, 4, 9, 17—22] розглянуто багато законів випадкового розподілу зв'язків — від простого рівномірного розподілу до експоненціального сімейства розподілів ймовірностей. Однією з перших базових моделей випадкового графа є модель Ердоса–Реніі [20], що має фіксовану кількість вершин та рівномірний розподіл зв'язків. За цих припущень ймовірнісний розподіл ступеня вершини визначається так:

$$P(d(n_i) = k) = \frac{(g-1)!}{k!(g-1-k)!} p^k (1-p)^{g-1-k}, \quad (14)$$

де p — ймовірність існування зв'язку між вершинами.

Розподіл (14) наближається до Пуассонівського коли $g \rightarrow \infty$. Якщо для всіх зв'язків $p = \frac{1}{2}$, отримаємо вираз:

$$P(d(n_i) = k) = \frac{(g-1)!}{k!(g-1-k)!} p^{g-1}. \quad (15)$$

Модель (15) дозволяє створити випадкову досліджувану мережу з наперед заданими властивостями, наприклад, кількістю ребер, та навіть оцінити розмір найбільшої компоненти графа [20].

Емпіричні дослідження [17—19, 23] свідчать про безмасштабний характер реальних мереж, створених людиною чи природою. Тобто, закон розподілу ступеня вершин є показовим. Для відображення безмасштабності мережі розроблена модель Барабаші–Альберта [17], що відноситься до еволюційних динамічних мереж. На кожному кроці до початкової мережі додається вершина з кількома зв'язками. При цьому, більше шансів утворити новий зв'язок має вершина, що уже має найбільше зв'язків. Такий принцип росту мережі змістовно інтерпретується правилом «багатий стає ще багатшим». В результаті зростання отримуємо мережу, яка має окремі потужні вузли та розрізнену периферію, що часто спостерігається для реальних систем. Ймовірність додавання зв'язку X_{ii} на кроці t в моделі Барабаші–Альберта визначається через відношення ступенів вершини:

$$P(X_{ii} = 1) = \frac{d(n_i)}{\sum d(n_j)}. \quad (16)$$

Ще одним явищем реальних мереж є феномен «малого світу» — відносно невеликий діаметр графа порівняно з кількістю вершин. Так, середня довжина зв'язку між двома жителями США складає всього 6 вершин [23]. Для дослідження цього явища розроблена модель Воца–Строгаца [22], згідно з якою з g вершин створюється кільцевий граф. Кожна вершина графа може з однаковою ймовірністю p перевизначити свій зв'язок з іншими вершинами. За цією моделлю можна створити мережу з наперед заданим діаметром. Встановлено, що середня довжина зв'язку між двома будь-якими вершинами мережі за умови $p \rightarrow 1$, наближається до

$$L = \frac{\ln g}{\ln d}. \quad (17)$$

Статистичний аналіз широко застосовується для побудови алгебраїчних моделей [1], знаходження тенденції обопільності вибору або взаємності [21], для вивчення еволюції мережі з плином часу [9], порівняння багатоарних відношень [1] тощо. Статистичний підхід є одним з найпродуктивніших і перспективних напрямків аналізу соціальних мереж, на базі якого досліджують складні багато-параметричні моделі в маркетингу [24], стохастичні моделі для дводольних графів [25] та моделі прихованого простору [26].

Моделі інших типів

В останнє десятиріччя завдяки бурхливому розвитку інформаційних технологій, стало можливо збирати та обробляти інформацію про топологію реальних мереж з мільйонами вершин. Також вільний доступ до інформації сприяє міждисциплінарним дослідженням складних соціальних мереж за допомогою гібридних моделей.

Одним із очевидних напрямків гібридизації є графо-ігровий підхід, що поєднує теорію графів, яка вивчає структуру мережі, з теорією ігор, яка досліджує моделі поведінки за умов, коли рішення акторів впливають одне на одного. Графо-ігровий підхід застосовують для аналізу структури мережі ринків та влади в них, моделювання взаємодії між окремими особами і населенням в цілому, моделювання динаміки епідемій та процесу пошуку в інформаційних системах [5].

В контексті теорії управління [27] соціальні мережі аналізують з урахуванням таких понять, як агент, думка, впливовість, довіра, репутація, дифузія тощо. Застосовуються методи аналізу на основі порогових моделей, моделей незалежних каскадів, моделей просочування та зараження, моделей Ізінга, клітинних автоматів та ланцюгів Маркова.

Останнім часом стрімко поширюється застосування моделей статистичної фізики для аналізу соціальних мереж [3]. Стаття [17] містить детальний огляд застосувань моделей статистичної фі-

зики та механіки для аналізу соціальних мереж, наприклад, розглянуто зв'язок між теорією перколяції (просочування) та моделлю випадкового графа Ердоса–Реніі. Теорія перколяції описує поведінку пов'язаних кластерів у випадковому графі. З цієї теорії випливає, що якщо в моделі (14) ймовірність p стане нижче деякого критичного значення p_c , тоді мережа складатиметься з ізольованих кластерів, а у випадку $p > p_c$ — гігантська компонента охопить усю мережу.

Ще одна цікава модель статистичної фізики — це фітнес-модель Альберта–Барабаші [17]. Оригінальна модель Альберта–Барабаші (16) відображає безмасштабність реальних мереж, але не враховує можливість появи нової аттракторної вершини, з якою намагатимуться встановити зв'язки багато інших вершин. Для врахування цього у фітнес-модель додано параметр η_i , значення якого для нової вершини обирається за деяким ймовірнісним розподілом. За високого рівня фітнес-параметра нові вершини можуть мати інтенсивний приріст зв'язків. Фітнес-модель записується таким чином:

$$P(X_{ii} = 1) = \frac{\eta_i d(n_i)}{\sum \eta_j d(n_j)}. \quad (18)$$

Оскільки люди спілкуються та думають, використовуючи переважно вербальні конструкції, R. R. Yager в [28] запропонував поєднати строгі математичні моделі соціальних мереж з лінгвістичними оцінками характеристик їх елементів. Наприклад, описуючи такі концепти соціальної мереж як «лідер» або «законодавець моди» аналітик використовує лінгвістичні терми з відповідного словника, і цей вербальний опис слід узгодити зі строгими математичними моделями мереж. В роботі [28] запропоновано концепцію, за якою можна оброблювати лінгвістичні запити до соціальної мережі з використанням нечітких множин. Засобами нечіткої математики досить легко формалізувати словесні поняття, наприклад, для опису зв'язності підгрупи за допомогою термів «Дуже низька», «Низька», «Середня», «Висока» тощо. Для багатофакторної ідентифікації зв'язних підгруп можна побудувати базу нечітких правил «якщо — тоді», які будять враховувати різномірну експертно-експериментальну інформацію про структурні та композиційні змінні.

Особливості аналізу онлайн соціальних мереж

Усі різноманітні онлайн соціальні мережі можна описати кортежем з двох множин: множини користувачів та множини контенту (Р. Яворський) [29]. Соціальний граф користувачів є мультиграфом, тобто користувачі утворюють багатоарне відношення, що може містити додаткові структурні дані

$$G = \{V, E_1 \dots E_k, p, \delta_1 \dots \delta_k\}, \quad (19)$$

де V — актори; $E_1 \dots E_k$ — типи відношень у мережі; p — профіль актора $V \rightarrow p$; $\delta_1 \dots \delta_k$ — параметри відповідного відношення $E_i \rightarrow \delta_i$.

Аналогічно можна описати граф контенту мережі

$$C = \{T, R_1 \dots R_m, \theta, \gamma_1 \dots \gamma_m\}, \quad (20)$$

де T — набір різномірного контенту мережі; $R_1 \dots R_m$ — відношення контенту між собою; θ — параметри контенту $T \rightarrow \theta$; $\gamma_1 \dots \gamma_m$ — параметри відповідного відношення у середині контенту $R_j \rightarrow \gamma_j$.

І нарешті, через $A \in V \times T$ — відношення акторів і контенту

$$A = \{L_1 \dots L_n, \varepsilon_1 \dots \varepsilon_n\}, \quad (21)$$

де $L_1 \dots L_n$ — відношення акторів мережі та контенту; $\varepsilon_1 \dots \varepsilon_n$ — відповідні параметри відношення $L_l \rightarrow \varepsilon_l$.

Відношення між акторами мережі та контентом, можна розглядати через авторство повідомлення (post) та коментарі (comments). Повідомлення можуть бути у формі тексту, зображенням, аудіо чи відео. Вони сповіщають учасників мережі про деяку подію чи явище. Коментарі — це зафіксоване у певному форматі (рис. 2) ставлення акторів до повідомлень.

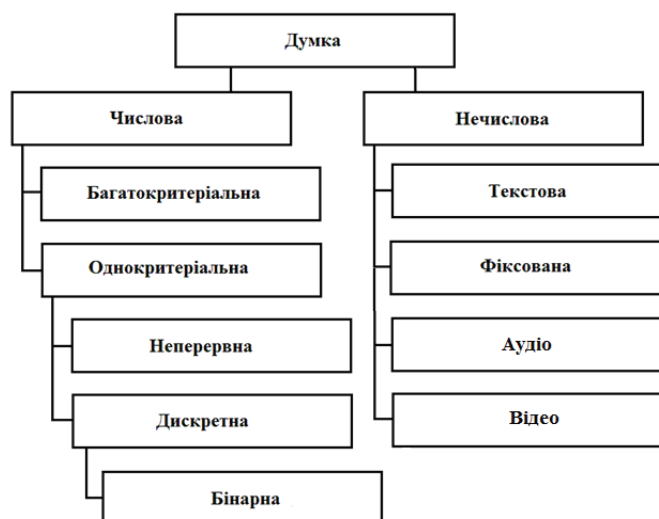


Рис. 2. Класифікація коментарів

Найпростіший аналіз онлайн соціальних мереж здійснюють на основі егоцентричного графа. Вершини цього графа відповідають центральному актору — «Его» та його найближчим сусідам. Такий граф не відображає характеристики мережі, але його можна застосувати для дослідження соціальних ролей у групі. Деякі шаблони можна виявити, аналізуючи топологію мережі. Але у більшості випадків потрібно знати такі характеристики мережі, як центральність, щільність, середня кількість шляхів що проходять через актора [14], рівень структурної еквівалентності з певним шаблоном. Інколи потрібні і статистичні характеристики: дисперсія, централізованість, гістограми розподілу ступеня вершини та інші. Шаблони відношень у егоцентричній мережі знаходять своє змістовне відображення у вигляді посади, соціальної чи професійної діяльності актора [30]. Наприклад менеджери та адміністратори поділяють між собою шаблон «містка», який зв'язує окремі підгрупи — структурні підрозділи підприємства. Шаблон «цибулина» добре підходить для опису команди розробників, де їх щільне ядро оточене шарами з іншим структурних підрозділів, з якими вони контактують [30]. Кожна соціальна роль формує свою структуру зв'язків в мережі, а отже може бути визначена з неї. Таким чином, за характером та структурою зв'язків можна класифікувати акторів, наприклад, виявити потенційно корисних користувачів та включити їх в систему контекстного пошуку, виявити «провокаторів», спамерів, «ботів», «флудерів» та відгородити їх від інших акторів.

Складніші дослідження проводять за повномасштабним графом мережі. Дослідження популярних онлайн соціальних мереж Flickr, YouTube, LiveJournal Orkut та Facebook свідчать про їх безмасштабний характер [1, 15, 18, 19, 23, 31, 32]. Це означає, що мережа має степеневий розподіл зв'язків $P(k) \sim k^{-\gamma}$ між акторами. Тобто, більшість акторів має однакову кількість зв'язків, а доля акторів з надмірно великою кількістю зв'язків — рідкість. Тому онлайн мережу можна представити досить розгалуженою структурою з ієрархічно пов'язаними вершинами на зразок комп'ютерних мереж. Безмасштабні мережі є відмовостійкими — вони майже не реагують на відмову випадкового елемента [15]. Водночас вони чутливі до акцентованих атак — цільове виключення елементів з найвищою кількістю зв'язків може призвести до втрати зв'язності компонент. Для розбиття соціальної мережі на малі компоненти достатньо вивести біля 1—10 % елементів з найвищим ступенем вершини [23]. Ці властивості є дуже важливими в аналізі соціальних мереж, так як розвиток теорії відмовостійкості та вразливості прямо пов'язаний з аналізом потоку інформації. Зв'язок між вузлами здійснюється переважно концентраторами — акторами з найбільшою кількістю зв'язків. Управління концентраторами майже тотожне підпорядкуванню усієї мережі, тоді як контроль великої кількості окремих акторів здебільшого є не виправданим.

В онлайн соціальних мережах спостерігається феномени безмасштабних мереж, наприклад, «багатий стає багатше» [32] та «світ тісний» [32, 23]. Але онлайн соціальні мережі мають низку відмінностей від інших безмасштабних мереж. Наприклад, на відміну від веб-мережі, в соціальній мережі переважають симетричні зв'язки [32, 23]. Внаслідок великої тісності та взаємності зв'язків онлайн соціальна мережа утворює тісніший кластер з діаметром біля 7. Діаметр веб-мережа стано-

вить біля 16, а малі групи утворюють повнозв'язні графи [23]. Є особливості у тенденціях створення нових зв'язків між акторами. Наприклад, користувачі мережі Facebook здебільшого не встановлюють зв'язки з людьми, яких вони не знають поза мережею [32]. Це виправдано через небезпеку шахрайства, оскільки в мережі Facebook близько 84 % користувачів не обмежують права доступу до своїх сторінок [31].

Збір та аналіз даних повної мережі вимагає неабияких матеріальних ресурсів. А. Mislove, M. Marcon, K. P. Gummadi [23] дослідили кілька соціальних мереж великого масштабу, зібравши публічно доступні зв'язки користувачів з чотирьох популярних соціальних онлайн мереж. Набір даних містить інформацію про понад 11,3 млн користувачів та 328 млн посилань. Це перше дослідження з вивчення кількох онлайн соціальних мереж у великому масштабі, що підтвердило безмасштабність мережі та низку феноменів соціальних мереж зазначених вище. Сьогодні кількість учасників соціальних мереж може нараховувати близько 10^8 , кількість зв'язків між ними досягає 10^{10} , і кожна з соціальна мережа, як і комп'ютерні мережі в цілому, являються безмасштабними мережами.

Щоб гарантовано отримати адекватні оцінки характеристик соціальних онлайн мереж необхідно розглянути повномасштабний граф. Методи отримання відповідних вибірок даних досить прості, а сам процес легко розпаралелити. Але сама вибірка може не виправдати матеріальні та обчислювальні витрати на її отримання. Тому для більшості дослідників отримання, а тим більше аналіз, повної мережі представляється неможливим. Для адекватної оцінки характеристик мережі достатньо отримати репрезентативну вибірку, яка має ті ж характеристики, що і повна мережа. Така мережа може бути зібрана шляхом випадкового вибору деякої підмножини акторів. Для отримання рівномірних вибірок існує декілька методів, найвідоміші з них [31]: обхід в ширину; випадковий обхід; повторно-зважений випадковий обхід; випадковий обхід Метрополіс–Хастінга.

Загалом обхід в ширину та випадковий обхід достатній для збору даних у необхідній кількості, але вибірка далека від рівномірної; її характеристики зміщені. Обхід Метрополіс–Хастінга та повторно-зважений випадковий обхід дають показники близькі до рівномірної вибірки з незміщеними характеристиками. Для знаходження статистично незміщених характеристик мережі достатньо розглянути підграф приблизно з 3...10 тисяч акторів [31].

Аналіз та відображення даних, зібраних з онлайн соціальних мереж, вимагає використання спеціалізованого програмного забезпечення, що включає пакети UCINET, Pajek, Gephi, ORA, NetMiner, StOCNET, MultiNet, GUESS, доповнення NodeXL та бібліотеки igraph (R, Python, C), libSNA, NetworkX (Python), SNA (R), SNAP(Gauss), SNAP (C++). Вичерпну інформацію про це програмне забезпечення можна отримати з [9, 33] та з відповідних сайтів виробників. Для обробки та зберігання даних профілі користувачів структурують. Для цього з веб-сторінки облікового запису в першу чергу виокремлюють ознаки і атрибути. Необхідно також визначити перелік ключових слів та провести їх нормалізацію, використовуючи для створення текстового індексу, наприклад, підхід на основі моделі мішка слів (bag of words) та стеммінгу [34].

У моделі мішку слів текст представляється невпорядкованим набором слів, без урахування граматики і порядку слів. Далі, для виділеної з профілю користувача інформації, необхідно провести стемінг — знайти основи заданого слова. Одним з найбільш популярних і ефективних алгоритмів стемінга є стемер Портера, що полягає у вилученні префіксів та суфіксів. Але цей метод має такі недоліки: 1) в українській та російських мовах, де відбувається чергування голосних та приголосних, слово зазвичай обрізується до занадто короткої основи; 2) метод чутливий до орфографічних помилок; 3) набір словотворчих частин різний для різних мов, як власне і словотворчі правила. Альтернативою стемінгу є нечіткий пошук слів, який базується на знаходженні набору слів в словнику за допомогою метрики Левенштейна. Ця метрика визначається за мінімальною кількістю виправлень, які необхідно внести для отримання слова зі словника. Перевагами цього методу є стійкість до помилок, мово-незалежність та степеневе порівняння схожості слів. Концептуальний підхід такої інформаційної технології аналізу соціальних мереж окреслений в роботі [35].

В онлайн соціальних мережах достатньо легко можна провести різні опитування та анкетування з відповідним аналізом за допомогою методів соціометрії. Отримана від користувачів соціальних мережі інформація містить різні види невизначеностей через дію різноманітних «Нефакторів». Для врахування невизначеностей початкової інформації для аналізу соціальних мереж в [36] пропонується використання нечіткої логіки.

Висновки

Початковий етап досліджень з аналізу соціальних мереж полягав у побудові графів на основі соціальної мережі та обчислення таких характеристик графа, як ступінь зв'язності, густина зв'язків тощо. За кількістю та направленістю зв'язків визначається центральність та престиж актора у спостережуваній групі. Перші методи кластеризації акторів базувались на теорії збалансованості та транзитивності зв'язків. Недоліки цих методів зумовлені тим, що збалансованих чи транзитивних мереж у світі не існує. Крім того, кластеризація на основі міри збалансованості чи транзитивності може привести до розбиття, яке суперечить апріорним відомостям.

Другий етап досліджень з аналізу соціальних мереж стосувався розробки теорії структурної еквівалентності та її узагальнення. Структурна еквівалентність надала поняття позиції актора, що змістовно інтерпретується відповідно до предметної області, наприклад, позиція на службі чи в ієрархії. Об'єднання у блок усіх акторів з однаковими позиціями, дозволяє укрупнити мережу, що значно спростило її дослідження. В реальних мережах немає повністю структурно еквівалентних акторів, тому постає проблема не тільки віднесення актора до певного блоку, але й визначення умов існування зв'язку між групами.

Одним з останніх етапів розвитку теорії аналізу соціальних мереж стали статистичні методи, в яких зв'язки між акторами описуються випадковими величинами. Статистичні моделі дають змогу визначати схильність акторів створювати нові зв'язки, створювати моделі підгруп та оцінювати невідомі атрибути акторів. Часто використовуються моделі статистичної фізики з теорії просочування та статистичної термодинаміки.

Обмеження традиційних методів ідентифікації залежностей зумовлені вимогами статистичної незалежності спостережень, які для реальних соціальних мереж ніколи не виконуються. Крім того, велика складність реальних мереж та їх варіативність катастрофічно збільшують алгоритмічну складність моделювання. Також, за традиційних підходів важко коректно врахувати такі особливості соціальних мереж як лінгвістичну невизначеність у поведінці акторів. Внаслідок цього перспективним інструментом для розв'язання задач ідентифікації залежностей в соціальних мережах виступає апарат нечіткої логіки. Ідея полягає в доповненні традиційних підходів до ідентифікації в соціальних мережах технологіями нечіткої логіки. Для формалізації поведінки акторів перспективним є застосування принципів ройового інтелекту, які за самим своїм походженням відображують основні закономірності функціонування людини в соціальних мережах. Означені інтелектуальні технології добре зарекомендували себе під час розв'язання задач моделювання складних систем з аналогічними особливостями.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Wasserman S. *Social Network Analysis: Methods and Applications* / S. Wasserman, K. Faust. — Cambridge: Cambridge University Press, 1994 — 857 p.
2. Martino F. *Social Network Analysis: A brief theoretical review and further perspectives in the study of Information Technology* / F. Martino, A. Spoto // *Psychology Journal*. — 2006. — Vol. 4, No 1. — P. 53—86.
3. Scott J. *Social network analysis developments, advances, and prospects* / J. Scott // *Social Network Analysis and Mining*. — 2011. — Vol. 1, Issue 1. — P. — 26.
4. Butts C. T. *Social network analysis. A methodological introduction* // C. T. Butts // *Asian Journal of Social Psychology*. — 2008. — Vol. 11. — P. 13—41.
5. Easley D. *Networks, Crowds, and Markets : Reasoning about a Highly Connected World* / D. Easley, J. Kleinberg. — Cambridge University Press, 2010. — 727 p.
6. Lorrain F. *Structural equivalence of individuals in social networks* / F. Lorrain, H. C. White // *The Journal of Mathematical Sociology*. — 1971. — Vol. 1, No. 1. — P. 49—80.
7. White H. C. *Social structure from multiple networks. I. Blockmodels of roles and positions* / H. C. White, A. S. Boorman, R. L. Breiger // *American journal of sociology*. — 1976. — Vol. 81, No. 4. — P. 730—780.
8. Borgatti S. P. *Notions of Position in Social Network Analysis* / S. P. Borgatti, M. G. Everett // *Sociological Methodology*. — 1992. — Vol. 22. — P. 1—35.
9. Carrington J. *Models and Methods in Social Network Analysis* / J. Carrington, S. Wasserman, G. Roberts. — Cambridge: Cambridge University Press, 2005 — 345 p.
10. Aral S. *The Diversity-Bandwidth Trade-off* / S. Aral, M. Alstynne // *American Journal of Sociology*. — 2011. — Vol. 117, No. 1. — P. 90—171.
11. Burt R. S. *Structural Holes and Good Ideas* / R. S. Burt // *The American Journal of Sociology*. — 2004. — Vol. 110, No. 2. — P. 349—399.
12. Borgatti S. P. *LS Sets, lambda sets and other cohesive subsets* / S. P. Borgatti, M. G. Everett, P. R. Shirey // *Social Network*. — 1990. — Vol. 12. — P. 337—357.

13. Burt R. S. Some properties of structural equivalence measures derived from sociometric choice data / R. S Burt // *Social Networks*. — 1988. — Vol. 10. — P. 1—28.
14. Borgatti S. P. Identifying sets of key players in a social network / S. P. Borgatti // *Computational & Mathematical Organization Theory*. — 2006. — Vol. 12. — P. 21—34.
15. Albert R. Error and attack tolerance of complex networks / R. Albert, H. Jeong, A.-L. Barabási // *Nature*. — 2000. — Vol. 406. — P. 378—382.
16. Breiger R. L. The Analysis of Social Networks / R. L. Breiger, edited by M. Hardy, A. Bryman. // *Handbook of Data Analysis*. — London : SAGE Publication., 2004. — P. 505—526.
17. Albert R. Statistical mechanics of complex networks / R. Albert, A.-L. Barabási // *Reviews of modern physics*. — 2002. — Vol. 74. — P. 47—97.
18. Barabasi A. L. Emergence of Scaling in Random Networks / A. L. Barabasi, R. Albert // *Science*. — 1999. — Vol. 286. — P. 509—512.
19. Clauset A. Power-law distributions in empirical data / A. Clauset, C. R. Shalizi, M. E. J. Newman // *SIAM Review*. — 2009. — Vol. 51, No. 4. — P. 661—703.
20. Newman M. E. J. The structure and function of complex networks / M. E. J. Newman // *SIAM review*. — 2003. — Vol. 45, No. 2 — P. 167—256.
21. Snijders T. A. B. Statistical Methods for Network Dynamics / T. A. B. Snijders // *Proceedings of the XLIII Scientific Meeting, Italian Statistical Society*. — Padova : CLEUP., 2006. — P. 281—296.
22. Watts D. J. Collective dynamics of 'small-world' networks / D. J. Watts, S. H. Strogatz // *Nature*. — 1998. — Vol. 393. — P. 440—442.
23. Mislove A. Measurement and Analysis of Online Social Networks / A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, B. Bhattacharjee// *IMC '07 Proceedings of the 7th ACM SIGCOMM conference on Internet measurement* — ACM New York, NY, USA., 2007. — P. 29—42.
24. Domingos P. Mining the Network Value of Customers / P. Domingos, M. Richardson // *KDD '01 Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*. — ACM New-York, NY, USA., 2001. — P. 57—66.
25. Zweig K. A. Kaufmann M. A systematic approach to the one-mode projection of bipartite graphs / K. A. Zweig, M. Kaufmann // *Social Network Analysis and Mining*. — 2011. — Vol. 1, Issue 1. — P. 187—218.
26. Hoff P. D. Latent Space Approaches to Social Network Analysis / P. D. Hoff, A. E. Rafterya, M. S. Handcocka // *Journal of the American Statistical Association*. — 2002. — Vol. 97, Issue 460. — P. 1090—1098.
27. Губанов Д. А. Социальные сети: модели информационного влияния, управления и противоборства / Д. А. Губанов, Д. А. Новиков, А. Г. Чхартишвили. — М. : изд-во физико-математической литературы, 2010. — 228 с.
28. Yager R. R. Fuzzy Soft Social Network Modeling and Marketing / R. R. Yager, R. L. Yager // *Fuzzy Methods for Customer Relationship Management and Marketing: Applications and Classifications*. — IGI Global Snippet. 2012. — P. 16—40.
29. Yavorsky R. Research Challenges of Dynamic Socio-Semantic Networks / R. Yavorsky // *CEUR Workshop proceedings*. — 2011. — Vol. 757. — P. 119—122.
30. Fisher D. Using egocentric networks to understand communication / D. Fisher // *Internet Computing, IEEE*. — 2005 — Vol. 9, No. 5. — P. 20—28.
31. Practical Recommendations on Crawling Online Social Networks / M. Gjoka, M. Kurant, C. T. Butts, A. Markopoulou // *IEEE Journal on Selected Areas in Communications*. — 2011. — Vol. 29, No. 9. — P. 1872—1892.
32. Tastes, ties, and time A new social network dataset using Facebook.com / [K. Lewis, J. Kaufman, M. Gonzalez, A. Wimmer, N. Christakis] // *Social Networks*. — 2008. — Vol. 30. — P. 330—342.
33. Cobo M. J. Science Mapping Software Tools Review, Analysis, and Cooperative Study Among Tools / M. J. Cobo, A. G. López-Herrera, E. Herrera-Viedma, and F. Herrera // *Journal of the American Society for Information Science and Technology*. — 2011. — Vol. 62, No. 7. — P. 1382—1402.
34. Куликова А. А. Подход к классификации пользователей социальных сетей / А. А. Куликова // *Восточно-европейский журнал передовых технологий*. — 2011. — Т. 3, № 2. — С. 14—18.
35. Губанов Д. А. Концептуальный подход к анализу онлайн-социальных сетей / Д. А. Губанов, А. Г. Чхартишвили // *Управление большими системами*. — 2013. — № 45. — С. 222—236.
36. Кондратенко Н. Р. Нечеткие модели принятия решений в задачах прогнозирования взаимоотношений в социальных группах / Н. Р. Кондратенко, С. В. Лужецкий // *Наукові праці ВНТУ*. — 2009 — № 2.

Рекомендована кафедрою комп'ютерних систем управління ВНТУ

Стаття надійшла до редакції 30.10.2014

Мазуренко Віктор Володимирович — аспірант кафедри комп'ютерних систем управління, e-mail: viktor.mazurenko@gmail.com;

Штовба Сергій Дмитрович — д-р техн. наук, професор, професор кафедри комп'ютерних систем управління

V. V. Mazurenko¹
S. D. Shtovba¹

Overview of models for social network analysis

¹Vinnytsia National Technical University

The history of social network analysis is considered in the paper. The main types of social networks and methods for their analysis are examined in detail. The relevance and application of methodology for the analysis of social networks, especially in today information technology are revealed in the article.

Keywords: social networks, social network analysis, hyperlinks network, quantitative indicators of social networks, block models, random graphs.

Mazurenko Viktor V. — Post-Graduate Student of the Chair of Computer Control Systems, e-mail: viktor.mazurenko@gmail.com;

Shtovba Serhii D. — Dr. Sc. (Eng.), Professor, Professor of the Chair of Computer Control Systems

В. В. Мазуренко¹
С. Д. Штовба¹

Обзор моделей анализа социальных сетей

¹Винницкий национальный технический университет

Рассмотрена история развития анализа социальных сетей. Подробно рассмотрены основные типы социальных сетей и методы их анализа. Раскрыта актуальность развития и применения методологии анализа социальных сетей, в частности в современных информационных технологиях.

Ключевые слова: социальные сети, анализ социальных сетей, сети гиперссылок, количественные показатели социальных сетей, блочные модели, случайные графы.

Мазуренко Виктор Владимирович — аспирант кафедры компьютерных систем управления, e-mail: viktor.mazurenko@gmail.com;

Штовба Сергей Дмитриевич — д-р техн. наук, профессор, профессор кафедры компьютерных систем управления