УДК 67.401.212.73
DOI: 10.31673/2409-7292.2019.034250

В.М. Ахрамович

МОДЕЛЬ ВЗАЄМОВІДНОСИН КОРИСТУВАЧІВ В СОЦІАЛЬНИХ МЕРЕЖАХ

Розглянуті елементи взаємовідносин користувачів в соціальних мережах: показано, що соціальне відношення користувачів мережі являє собою таблицю зв'язності; перетин матриць соціальних відносин та зв'язності; композицію відношень: елементи та типи графів мережі; набір статистики для соціальних відносин; інтенсивність взаємодії груп, ступінь центральності різних користувачів; характеристики мережі соціальних взаємодій - збалансованість і транзитивність; сила структурної позиції користувача; стани діад; вплив структури мережі на модель **р**; ймовірність існування зв'язку між користувачами; залежні та незалежні ребра графа; оцінка параметрів моделі; логіт-моделі.

Ключові слова: соціальна мережа, користувачі, відносини, групи, перетин матриць, композиція.

Вступ

З шаленим розвитком соціальних Інтернет-мереж дедалі частіше постає питання щодо моделювання останніх засобами формальної теорії, яка дала б змогу досліджувати явища та процеси ще до їх появи у реальному світі. Існує низка класичних теорій, започаткованих ще задовго до появи Інтернету та соціальних сервісів, але актуальність яких не втрачено через можливість їх застосування у теперішньому світі. Тому, у сучасній науці продовжуютья спроби адаптації відомих підходів для дослідження соціальних явищ у віртуальному середовищі.

Аналіз літературних джерел

Т. Ньюкомб [9] запропонував враховувати відносини, які встановлюються між агентами спілкування і між ними і об'єктом мови. Його модель отримала популярність під назвою ABX model (рис. 1). Вона стверджує, що хтось (A) посилає інформацію комусь іншому (B) про щось (X). Модель передбачає, що орієнтації A до B і до X залежать один від одного. В результаті формуються чотири орієнтації (рис. 1):

1. A to X orientation;
2. A to B orientation;
3. B to X orientation;
4. B to A orientation.

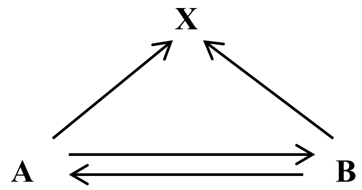

Рис. 1. Трикутник комунікації Т. Ньюкомба, Харари (Cartwright, Harary)

Якщо A і B зорієнтовані один до одного позитивно, то вони будуть прагнути до збігу свого ставлення до X. При розбіжності ставлення один до одного буде не збігатися і ставлення до X. Збіг відношення до X при розбіжності ставлення один до одного буде сприйматися як ненормальне.

Т. Нькомб стверджував, що існує тенденція, згідно з якою, якщо два індивіда близькі один до одного за поглядами, найімовірніше, вони займуть узгоджену позицію і до будь-якого третього індивіду, предмету або події. Спираючись на подібні відкриття, дослідники могли будувати моделі систематичної взаємозалежності між установками, яких дотримувалися різні індивіди в рамках однієї групи. Дане твердження було узагальнено в теоретичній концепції Картрайта і Харари (Cartwright, Harary) [1,2] . Спроба застосувати математику до структури групових відносин була, звичайно, не новою ідеєю. Такі спроби робилися і в кінці 1940-х років (Glanzer , Glaser. 1959: 317–332; Bavelas 1948: 16–30; 1950:

271–282; Festinger 1949: 153–158). Грунтуючись на роботі Левіна [7], Картрайт, Зандер і Харари (Cartwright, Harary 1956: 277–293) [1] розвинули потужні моделі групової згуртованості, соціального тиску, співпраці, влади і лідерства

У Великобританії для вивчення соціальної структури Редкліфф-Браун [19] застосовував два методи - морфологічний і фізіологічний вивчення соціальних систем. У функції першого входять визначення, порівняння та класифікацію різних структур. Завдання фізіологічного методу - вивчення механізмів, що підтримують існування системи

«Соціальна фізіологія ... - підкреслює РедкліффБраун, - має справу не тільки з соціальними структурами, але з усіма видами соціальних явищ.

У 1960-х роках був зроблений остаточний прорив до добре розвиненою методології аналізу соціальної мережі, і ця подія відбулася ще в Гарварді. Харрісон Уайт [3] поширив математику на дослідження соціальної структури, а його учні познайомили з ідеями свого вчителя весь світ.

Фриц Хайдер [4,5] вважається автором теорії когнітивного балансу - мотиваційної теорії зміни установок. У ній концептуалізується когнітивна узгодженість мотивів як причини психологічного балансу. Консистенція мотивів прагне підтримувати чиїсь цінності і переконання з плином часу. Хайдер припустив, що «настрої» або відносини збалансовані, якщо їх вплив примножує позитивний результат системи.

Якщо людина А любить Б, а людина В любить С. То стан балансу існує тільки тоді, коли А також любить С. Всі відносини є «позитивними». Важливо відзначити, що для Хайдера, як і для Левіна, цей вид аналізу ставиться до того, як світ сприймається з точки зору фокусної особистості. Хайдер стояв на феноменологічної позиції. З цієї точки зору головне не фактична зв'язок між В і С, а сприйняття точніше їх відносини.

В даний час вчені, використовуючи в своїх роботах мережеві підходи, проводять дослідження із найрізноманітніших напрямів у вивченні соціальних проблем сучасного суспільства. У роботі В. Печенкина і Е. Ярською-Смирнової [18] викладені результати емпіричного аналізу структурних особливостей функціонування віртуальних спільнот в соціальних мережах, який заснований на використанні різних методологічних підходів до дослідження. Представлено опис мікс-стратегії, способу спільного використання стратегій якісного і кількісного дослідження. Джерелом даних для аналізу соціальної згуртованості віртуальних спільнот є соціальна мережа.

К.Є. Гурин [14] розглядає структуру дружніх зв'язків учасників офіційних онлайн-спільнот російських ЗМІ в соціальній мережі «ВКонтакте», визначаючи онлайн-спільноти ЗМІ як новий феномен, що з'явився в результаті проникнення традиційних мас-медіа на платформи сучасних соціальних онлайн-мереж і концентрації в цих спільнотах аудиторії. Дослідницький інтерес сфокусований на процесах формування аудиторії, її структури і взаємодіях, а також на фактори впливу на процеси формування аудиторії. Емпірично обгрунтовується актуальність вивчення мереж онлайн-дружби в спільнотах

На думку А.В. Назарчука, [17] що досліджував витоки методу мережевого аналізу в соціальній сфері, теоретичні постулати мережевого аналізу, і мережеві підходи в рамках досліджень груп ці уявлення, дійсно плідними можуть стати тільки в поєднанні з сильною теоретичною наукою, здатної надати мережевим дослідженням, не тільки практичний, емпіричний або маркетинговий результат, а й масштаб макросоціологічного дослідження.

Цікавою є роботи [15,16] Д.В. Мальцевої, в якій системно викладено теоретичні ідеї та практичні напрацювання мережевого підходу в соціології та проведено аналіз соціальних мереж, роль мережевого підходу в структурі соціологічного знання. В роботі проаналізовано теоретико-методологічні основи трьох напрямків мережевого підходу в соціології - аналізу соціальних мереж, реляційної соціології та акторно-мережевий теорії. Аналіз зарубіжних і українських досліджень, присвячених мережевого підходу і його напрямками, дозволив скласти інформаційну базу дослідження, але в той же час привів до висновку про досить скромному уявленні даної проблематики у вітчизняній літературі.

Важливою складовою державних інформаційних ресурсів є персональні дані (ПД). Сьогодні забезпечення захисту інформації потребує неабиякої уваги, про що наголошується у рішенні Ради національної безпеки і оборони України «Про заходи щодо вдосконалення формування та реалізації державної політики у сфері інформаційної безпеки України» від 28 квітня 2014 [1]. А саме про необхідність вживання додаткових заходів щодо захисту інформації з обмеженим доступом (насамперед ПД, що належать до конфіденційної інформації) під час її обробки в автоматизованих системах (АС). Згідно з вимогами Закону України «Про захист інформації в інформаційно-телекомунікаційних системах» для забезпечення безпеки державних інформаційних ресурсів, оброблюваних в автоматизованій системі, необхідно розробляти комплексну систему захисту інформації (КСЗІ).

Взаємовідносини користувачів в соціальних мережах визначають, в тому числі, стан стійкої життєдіяльності мережі,. проблеми захисту їх персональних даних, реалізації основних інтересів і пріоритетних цілей користувача, розкриття маршрутів та колег по спілкуванню, зміст даних, що пересилаються, вподобання та переваги, які надають користувачі різного типу інформації, захист від зовнішніх і внутрішніх дестабілізуючих факторів незалежно від умов функціонування і т.п.

Аналіз проблем конфіденційності в соціальних мережах (СМ) демонструє, що навіть якби всі учасники були обізнані та компетентні у використанні Інтернет соціальних мереж (ОСМ), і навіть якщо було розгорнуто набір заходів щодо конфіденційності, ОСМ все од-но піддаватиметься потенційним порушенням конфіденційності будь-яким всезнаючим постачальником послуг або зовнішнім зловмисником, який бере під контроль ОСМ.

Тому у процесі розроблення концепції моделей взаємовідносин користувачів в соціальних мережах варто виділити основні процеси взаємодії і виключити можливість витоку інформації, її несанкціонованого використання, нанесення збитків, упущення вигоди з боку всіх зацікавлених сторін і в напрямі досягнення основних цілей захисту. Реалізація цих положень гармонійно вписується в концепцію управління соціальними мережами, до якої сьогодні залучаються дедалі більше користувачів.

Основна частина

Мережа соціальних взаємодій складається із сукупності соціальних користувачів і набору зв'язків між ними. Як соціальні користувачів можуть виступати індивіди, соціальні групи, організації, міста, країни. Під зв'язками розуміються не тільки комунікаційні взаємодії між користувачами, а й зв'язки з обміну різними ресурсами і діяльністю, включаючи конфліктні відносини.

Визначимо соціальну мережу як набір з g соціальних користувачів і r соціальних відносин, що показують як користувачі взаємодіють один з одним. Позначимо безліч користувачів як N={1, 2, ..., g}, а сукупність соціальних відносин як R={1, ... , r}. соціальне відношення χ як безліч впорядкованих пар номерів користувачів виду (i, j), де на першому місці стоїть номер вершини, з якої виходить дане ребро, а на другому місці - номер вершини, до якої входить це ребро. Більш наочно таку пару можна позначити як $i \rightarrow j$. Соціальне відношення може бути як направленим (незалежно один від одного можуть існувати ребра $i \rightarrow j$ та $j \rightarrow i$, так і ненаправленим (одне ненаправлене ребро з'єднує користувачів s та j). При цьому ребра можуть мати різні ваги, що показують силу взаємодії. Кожне соціальне відношення являє собою таблицю зв'язності X розміру $g \times g$, в якій елемент

$$x_v = \begin{Bmatrix} 1, \text{якщо існує ребро } i \rightarrow j \\ 0, \text{в протилежному випадку} \end{Bmatrix}$$

В даному випадку ми маємо r соціальних відносин $\chi_1, ..., \chi_r$. Позначимо відповідні їм матриці зв'язності як $X_1, ..., X_r$. Можна розглядати перетин $\chi_m \cap \chi_n$ відносин χ_m і χ_n, що задається матрицею $X_m \cap X_n$:

$$(X_m \cap X_n)_{Uij} = \begin{cases} 1, & \text{якщо } (X_m)_{ij}=1 \text{ і } (X_n)_{ij}=1 \\ 0, & \text{в протилежному випадку} \end{cases}$$

Композиція відношень $\chi_m\chi_n$ задається матрицею X_mX_n:

$$(X_m X_n)_{ij} = \begin{cases} 1, & \text{якщо } (X_m)_{ij}=1 \text{ і } (X_n)_{ij}=1 \text{ для деякої вершини } l \\ 0, & \text{в протилежному випадку} \end{cases}$$

Визначимо спрямований шлях довжиною d з вершини i в вершину j як набір вершин $\{i=i_1, i_2, \ldots, i_d = j\}$, такий що $x_{i_1i_2}x_{i_2i_3}\ldots x_{i_{d-1}i_d} = 1$. Найкоротший шлях з вершини i в вершину j будемо називати дистанцією або відстанню між вказаними вершинами і позначимо його довжину як $d(x_i, x_j) = d_{ij}$. Відстань не визначена, якщо не існує шляху між даними вершинами. Будемо називати граф, в якому деякі відстані не визначені, незв'язним графом або графом, що складається з декількох компонентів.

Для кожної соціальної відносини розглядається великий набір статистики. Зазначимо лише найбільш часто використовувані з них. Вхідна ступінь вершини $D_{in}(i) = \sum_j x_{ij}$. Вихідну ступінь вершини $D_{out}(i) = \sum_i x_{ij}$. Число ребер $L = \sum_{i,j} x_{ij}$. Число симетричних (взаємних) діад, тобто діад, де одночасно існують ребра i→j и j→i: $M = \sum_{i<j} x_{ij}x_{ji}$. Число вихідних зірок розміру 2: $S_O = \sum_{i\neq j\neq k} x_{ij}x_{ik}$. Число вхідних зірок разміру 2: $S_I = \sum_{i\neq j\neq k} x_{ji}x_{ki}$. Число змішаних зірок розміру 2: $S_M = \sum_{i\neq j\neq k} x_{ji}x_{ik}$. Число циклічних тріад, тобто таких трех ребер, що i→j, j→k, k→i: $T_C = \sum_{i\neq j\neq k} x_{ij}x_{jk}x_{ki}$. Число транзитивних тріад: $T_T = \sum_{i\neq j\neq k} x_{ij}x_{jk}x_{ik}$. Число не транзитивних тріад: $T_I = \sum_{i\neq j\neq k} x_{ij}x_{jk}(1-x_{ik})$.

Вводимо також ряд статистик, які описують гомогенні еффекти - число шляхів довжиною k; середня відстань (близкість); мінімальне число ребер, віддалення яких розбиває граф на декілька частин (зв'язність); і індивідуальні властивості користувачів - середня відстань від вершини до решти вершин; число шляхів, що включають вершину та (проміжки).

Якщо користувачі розбиті на кілька груп, то визначимо індикаторну змінну:

$$\delta_{U,rs} = \begin{cases} 1, & \text{якщо користувач i знаходиться в групі } r \text{ та користувач } j \text{ знаходиться в групі } s. \\ 0, & \text{в протилежному випадку} \end{cases}$$

Тоді інтенсивність взаємодії груп r та s: $I^{rs} = \sum_{ij} x_{ij}\delta_{ij,rs}$ [1].

Для аналізу ступеня центральності різних користувачів використовується індекс центральності вершини x_i в зв'язаному графі G, що має одну вісь симетрії: $C(x_i) = \frac{S(x_i)}{S(G)}$, де $S(x_i) = \sum_{j=1}^{n} d(x_i, x_j)$ - сума відстаней від вершини x_i до інших вершин, $S(G) = \sum_{i=1}^{n} S(x_i)$ - загальна сума дистанцій в графі G [10].

Для порівняння графів за ступенем їх центральності вводиться індекс центральності графа G: $C(G) = \frac{\min_i S(x_i)}{S(G)}$. В [1] доказано, що виконується умова $\frac{1}{2(n-1)} \le C(G) \le \frac{1}{n}$, в силу якого C(G) можливо нормувати та записати в вигляді $\delta(G) = \frac{2(n-1)(1-nC(G))}{n-2}$. Коефіцієнт δ(G) = 0 в разі абсолютно не центрального графа (наприклад, замкнутого циклу) δ(G) = 1 для центрального графа (зірки).

Важливими характеристиками мережі соціальних взаємодій є збалансованість і транзитивність. Збалансованість - це відсутність ситуацій типу «позитивна взаємодія (дружба, партнерство) між А та Б, а також між А та В, але негативний взаємодія (ворожнеча, суперництво) між Б та В». Стверджується, що збалансовані мережі психологічно більш комфортабельні для користувачів і більш стійкі в порівнянні з незбалансованими [6]. Транзитивність - це виконання умов виду «якщо є взаємодія між А та Б, а також між Б та В, то має місце взаємодія між А та В». Дані характеристики описують локальні зв'язку користувачів і часто використовуються при аналізі діад і тріад.

Сила структурної позиції користувача. Основним показником, який визначає відмінності в ресурсах користувачів, є сила структурної позиції користувача. В теорії мережевого обміну для вимірювання даної характеристики вводиться індекс сили користувача i: $GPI_i = \sum_{k=1}^{g-1} (-1)^{k-1} P[i]_k$ де - $P[i]_k$ число шляхів, які не пересікаються, довжини k, що проходять через вершину i. Сила користувача i в порівнянні з j є $GPI_{ij} = GPI_i - GPI_j$ [3].

Аналіз сили структурних позицій є дуже ефективним методом побудови моделей, що пояснюють відмінності в результатах діяльності користувачів, наприклад, заробітної плати працівників або прибутку компаній від продажу деякого товару, особливо в разі малого числа спостережень.

Стохастичні моделі. Основна ідея ймовірнісних моделей спрямованих графів полягає в тому, що кожна соціальна мережа може бути розглянута як реалізація x = $\{x_{ij}\}$ випадкового двовимірного бінарного масиву X. Так як елементи масиву X є залежними випадковими величинами, то можна аналізувати структуру залежностей між відповідними користувачами соціальної мережі , знаходити ймовірності існування певних реалізацій соціальної мережі і отримувати оцінки її параметрів

Моделі p_1. Розглянемо діаду Dij = (xij; xji). Вона може знаходитися в одному з чотирьох станів: (0; 0) - нульова діада (немає взаємодії), (0; 1) і (1; 0) - асиметричні діади, (1; 1) - симетрична (взаємна) диада. Створимо нову матрицю Y розмірності g×g×2×2 за таким правилом:

$$Y_{ij,ki} = \left\{ \begin{matrix} 1, \text{ якщо діада } ij = (x_U = kx_{ij} = l \\ 0 \text{ ,в протилежному випадку} \end{matrix} \right\}.$$

Далі розглянемо наступну модель:

$$\log P(Y_{ij00} = 1) = \lambda_{ij} \quad (1а)$$

$$\log P(Y_{ij10} = 1) = \lambda_{ij} + \theta + \alpha_i + \beta_j \quad (1б)$$

$$\log P(Y_{ij01} = 1) = \lambda_{ij} + \theta + \alpha_j + \beta_i \quad (1в)$$

$$\log P(Y_{ij11} = 1) = \lambda_{ij} + 2\theta + \alpha_i + \alpha_j + \beta_j + \beta_i + \rho \quad (1г)$$

Параметр α описує схильність суб'єкта до встановлення взаємодії (його оцінкою є вихідна ступінь D_{out}), параметр β описує привабливість або популярність (оцінка - вхідна ступінь D_{in}), θ- щільність графа (оцінка - число ребер L), ρ- характеристика тенденцій моделі до симетричності діад (оцінка - число симетричних діад M). У разі ненапрямлених взаємодій є тільки два типи діад і співвідношення (1) будуть включати тільки два параметри: оцінку інтенсивності взаємодій та θ.

Припустимо, що взаємодії мають ваги в вигляді цілих чисел від 0 до С- 1 (С - деяка константа). В цьому випадку матриця Y матиме розмірність g×g×C×C і співвідношення (1г) набирає вигляду:

$$\log P(Y_{ijkl} = 1) = \lambda_{ij} + \theta_k + \theta_l + \alpha_{i(k)} + \alpha_{j(l)} + \beta_{j(k)} + \beta_{(l)i} + \rho_{kl}\,.$$

Незалежно від природи взаємодій діади передбачаються статистично незалежними і такі, що мають однаковий імовірнісний розподіл, отже, можна ввести деяку функцію правдоподібності, наприклад, добуток ймовірностей станів діад. Тому моделі типу p1 називають також моделями з незалежними діадами [3].

Моделі p *. Визначимо для дихотомічного спрямованого соціального відношення χ три нові матриці: $X^+_{ij} = \{x_{kl}, \textit{при } x_{ij} = 1\}$ - примусово додано ребро i→j, $X^-_{ij} = \{x_{kl}, \textit{при } x_{ij} = 0\}$ - примусово прибрано ребро i→j, X^C_{ij} - всі ребра, додаткові до i→j, причому самого цього ребра в матриці X^C_{ij} немає.

Припустимо, що ймовірність прийняття матрицею X значення x дорівнює: $P(\mathbf{X} = \mathbf{x}) = \dfrac{\exp(\boldsymbol{\theta}'\,\mathbf{z}(\mathbf{x}))}{k(\boldsymbol{\theta})}$, де **θ'** - транспонований вектор параметрів моделі, z (x) - вектор статистик соціальної мережі, k - лінійна функція, що підсумовує добуток елементів вектора **θ** і деяких коефіцієнтів, та забезпечує потрібний вид ймовірнісного розподілу. Основна проблема при такому формулюванні завдання полягає у визначенні k, що є досить складним завданням для більшості соціальних мереж. Однак, можна перетворити розглянуту логлінійну модель в логіт-модель, використовуючи те, що випадкова величина x_{ij} є дихотомічною.

Звідси ймовірність існування ребра i→j:

$$P(x_{ij} = 1 \mid X^C_{ij}) = \frac{P(X = \mathbf{x}^+_{ij})}{P(X = \mathbf{x}^+_{ij}) + P(X = \mathbf{x}^-_{ij})} = \frac{\exp(\boldsymbol{\theta}'\mathbf{z}(\mathbf{x}^+_{ij}))}{\exp(\boldsymbol{\theta}'\mathbf{z}(\mathbf{x}^+_{ij})) + \exp(\boldsymbol{\theta}'\mathbf{z}(\mathbf{x}^-_{ij}))}$$

$$\frac{P(x_{ij} = 1 \mid X^C_{ij})}{P(x_{ij} = 0 \mid X^C_{ij})} = \frac{\exp(\boldsymbol{\theta}'\mathbf{z}(\mathbf{x}^+_{ij}))}{\exp(\boldsymbol{\theta}'\mathbf{z}(\mathbf{x}^-_{ij}))} = \exp\{\boldsymbol{\theta}'[\mathbf{z}(\mathbf{x}^+_{ij}) - \mathbf{z}(\mathbf{x}^-_{ij})]\}\,. \quad (2)$$

Якщо позначити різницю в квадратних дужках як $\delta(x_{ij})$, то логарифм залежності (2) (логіт-модель) буде мати вигляд:

$$\omega_{ij} = \log\left(\frac{P(x_{ij} = 1 \mid X^C_{ij})}{P(x_{ij} = 0 \mid X^C_{ij})}\right) = \{\boldsymbol{\theta}'[\mathbf{z}(\mathbf{x}^+_{ij}) - \mathbf{z}(\mathbf{x}^-_{ij})]\} = \boldsymbol{\theta}'\delta(x_{ij})\,. \quad (3)$$

Елементи $\delta(x_{ij})$ - це зміни статистик соціальної мережі при зміні значення x_{ij} з 1 на 0. Такий варіант моделі, в якій логарифм відносини ймовірностей дорівнює лінійної комбінації елементів $\delta(x_{ij})$, називається логіт-моделлю p * для одиничного дихотомічного соціального відносини. Для випадку ребер з вагами від 0 до 1 ми будемо мати набір з С-1 логіт-моделей [12].

Статистична інтерпретація логістичних регресійних моделей для ω_{ij} залежить від припущення про незалежність величини ω_{ij}. У моделі p* логіт-моделі для ω_{ij} не є незалежною, в силу чого статистика відношення правдоподібності не піддається суворій статистичній інтерпретації, хоча її значення може служити орієнтиром якості побудованої моделі.

Графи залежності. На базі соціальної мережі можна побудувати граф залежності D, який показує, які ребра або групи ребер умовно залежні. Два ребра називаються умовно залежними, якщо умовна ймовірність одночасного існування цих ребер, що обчислюється за іншим ребрам мережі, не дорівнює добутку умовних ймовірностей незалежного існування цих ребер. Граф залежності має ребра, що зв'язують всі пари умовно залежних ребер соціальної мережі. Позначимо безліч вершин графа залежності як $N_D=\{(i,j,m);i,j\in N,i\neq j;m\in R\}$, (ми розглядаємо граф залежності для випадку декількох типів взаємодій).

За допомогою теореми Хаммерслі-Кліффорда можна формально встановити, як структура соціальної мережі впливає на параметри моделі p *. Дана теорема стверджує, що ймовірність існування випадкового спрямованого графа залежить тільки від повних підграфів в графі залежності і може бути представлена у вигляді:

$$P(\mathbf{X}=\mathbf{x})=\frac{1}{k}\exp\left(\sum_{A\subseteq N_D}\lambda_A\prod_{(i,j,m)\in A}\left(X_m\right)_{ij}\right),$$

де D - граф залежності для **X**, $k=\sum_{\mathbf{x}}\exp\left(\sum_{B\subseteq D}\lambda_B\prod_{(i,j,m)\in B}\left(X_m\right)_{ij}\right)$ – нормуюча константа, підсумовування йде по всіх підмножинах A множини ND, $\prod_{(i,j,m)\in A}\left(X_m\right)_{ij}$ – достатні статистики, відповідні параметру λ_A, причому $\lambda_A=0$ в разі, коли підграф, що задається вершинами, які входять в A, не є повним [13].

Ненульові параметри моделі відповідають набору максимальних повних підграфів (клік) графа залежності. Нагадаємо, що повним підграфом називається такий набір вершин, де кожна їхня пара пов'язана ребром, тобто повний підграф в графі залежності відповідає набору ребер соціальної мережі, кожна пара яких є умовно залежною. Максимальний повний підграф - це повний підграф, який в повному обсязі міститься в будь-якому іншому повному підграфі. Так як кожен підграф повного підграфа також є повним, тому, якщо A – це максимальна кліка D, то ненульові параметри моделі будуть при A і всіх його підграфах.

Використання теореми Хаммерслі-Кліффорда дозволяє істотно спростити процес побудови моделей p*. Наприклад, логлінійна модель Марківського графа залежить тільки від повного набору тріад і зірок розміру k, але не від тетрад і інших повних підграфів. При цьому модель можна додатково спростити за допомогою припущень про гомогенності її параметрів, тобто їх незалежності від індивідуальних користувачів.

Оцінка параметрів моделі p*. Функція правдоподібності для моделі p* може бути записана у вигляді $L(\theta)=\dfrac{\exp(\boldsymbol{\theta}'\mathbf{z}(\mathbf{x}))}{k(\boldsymbol{\theta})}$. Тоді знову виникає проблема визначення функції k. Уникнути цього можна при використанні так званої функції псевдоправдоподібності при припущенні умовної незалежності ребер і переході до логіт-моделі, яка має вигляд:

$$PL(\boldsymbol{\theta})=\prod_{i\neq j}\prod_{m=1}^{r}P((X_m)_{ij}=1\,|\,(\mathbf{X}_m^C)_{ij})^{(X_m)_{ij}}P((X_m)_{ij}=0\,|\,(\mathbf{X}_m^C)_{ij})^{1-(X_m)_{ij}} \quad (4)$$

Оцінка максимальної псевдоправдоподібності θ відповідає максимальному значенню виразу (4). Відзначимо, що оцінка максимальної правдоподібності не відрізняється від оцінки максимального псевдоправдоподібності тільки на класі найпростіших графових моделей, у яких умовні ймовірності існування ребер не залежить від структури ребер даної взаємодії, тобто в разі умовної незалежності ребер.

Д. Страуссом і М. Айкедой доведена теорема про те, що в разі логіт-моделі p * в формі (3) максимізація значення функції псевдоправдоподібності (4) еквівалентна максимізація функції правдоподібності для логістичної регресії в моделі (3) для незалежних спостережень $\{ x_{ij} \}$. Така апроксимація може бути виконана за допомогою ітеративного методу Гаусса-Ньютона з перезважуванням [4,5].

Позначимо оцінки параметрів моделі, отримані за допомогою логістичної регресії, як $(\tilde{X}_m)_{ij}$. Оцінювати якість апроксимації будемо за допомогою статистики $G_{PL}^2 = 2\sum (X_m)_{ij} \log \frac{(X_m)_{ij}}{(\tilde{X}_m)_{ij}}$. Для перевірки статистичної значущості кожної характеристики мережі, наприклад, взаємності, розглядаються дві моделі, що містить дану характеристику, і та, що не містить її. Значимість відмінностей в значеннях статистики G_{PL}^2 для цих моделей може бути наближено оцінена за допомогою розподілу χ^2 з числом ступенів свободи, що дорівнює кількості параметрів моделі, пов'язаним з цією характеристикою. Також значимість коефіцієнтів регресії можна перевірити за допомогою статистики Вальда, використовуючи їх наближену стандартну помилку.

Висновки

Сьогодні успішне функціонування соціальних мереж неможливо без чіткої і організованої системи управління інформаційною безпекою, яка відповідає багатьом стандартам і правовим актам. Необхідність захисту соціальних мереж обумовлена, в тому числі, станом стійкої життєдіяльності мережі, проблемами захисту персональних даних та реалізації основних інтересів і пріоритетних цілей користувачів, розкриття маршрутів та колег по спілкуванню, змісту даних, що пересилаються, вподобання та переваги, які надають користувачі різного типу інформації, захист від зовнішніх і внутрішніх дестабілізуючих факторів незалежно від умов функціонування і т.п.

Концепція захисту мереж та їх компонентів потребує від науковців створення моделей захисту, в тому числі і моделей взаємовідносин користувачів в соціальних мережах з метою подальшої розробки комплексних систем захисту.

Перелік посилань

1. Cartwright, D., & Harary, F. (1956). Structural balance: a generalization of Heider's theory. Psychological Review, 63(5), 277–293. https:// https://doi.org/10.1037/
2. Harary F., Norman R.Z. (1953) The dissimilarity characteristic of Husimi trees. Ann. of Math, 58: 134–141
3. Harrison C. White (1958). Varieties of Markets, in Barry Wellman and S.D. Berkowitz, (eds.), Social Structures: A Network Approach. Cambridge: Cambridge University Press.
4. Heider F. (1946) Attitudes and Cognitive Organization. Th e Journal of Psychology, 21: 107– 112.
5. Heider F. (1958) Th e Psychology of Interpersonal Relations. New York: John Wiley & Sons.
6. Johnson J., Ironsmith M. Assessing children's sociometric status: Issues and the application of social network analysis // Journal of Group Psychotherapy, Psychodrama & Sociometry, Spring94. Vol. 47. Issue 1. P. 36-49.
7. Lewin, K., and Dorwin Cartwright (Ed.) (1951). Field theory in social science. New York: Harper.
8. Lewin, K., and Gertrude W. Lewin (Ed.) (1948). Resolving social conflicts: selected papers on group dynamics [1935-1946]. New York: Harper and Brothers.
9. Newcomb T.M. An approach to the study of communicative acts // Psychol. Rew., 1953, v. 60, p. 293–304.
10. Parlebas P. Centralite et compacite d’un graphe // Mathematiques et sciences humaines, 1972. Vol. 39. P. 5-26.
11. Wasserman S., Iacobucci D. Statistical analysis of discrete relational data // British Jornal of Mathematical and Statistical Psychology, 1986. Vol. 39. P. 41-64.

12. Wasserman S., Pattison P. Logit models and logistic regressions for social networks: I. An introduction to markov graphs and p* // Psychometrika, 1996. Vol. 61. P. 401-425.

13. Wasserman S., Pattison P. Logit models and logistic regressions for social networks: II. Multivariate relations // British Journal of Mathematical and Statistical Psychology, 1999. Vol. 52. P. 169-193.

14. Гурин К.Е. (2016) Структурирование сетей дружбы в онлайн-сообществах СМИ. Дискуссия. 6 (69): 64–71.

15. Мальцева Д. В. (2017) Сетевой подход в социологии: генезис идей и применение. Новосибирск: НГТУ.

16. Мальцева Д. В. Сетевой подход как феномен социологической теории // Социологические исследования. 2018. № 4. С. 3-14

17. Назарчук А.В. (2011) О сетевых исследованиях в социальных науках. Социологические исследования. 1: 39–51.

18. Печенкин В., Ярская-Смирнова Е. (2014) Сетевые подходы в анализе социальной сплочённости. Вестник СГТУ, 4 (77): 249–253.

19. Рэдклифф-Браун Альфред. О понятии «функция» в социальных науках // У кн.: Рэдклифф-Браун Альфред. Структура и функция в примитивном обществе. Очерки и лекции / Пер. с англ. О. Ю. Артемовой при участии Ю. А. Артемовой, А. Г. Артемова. — Москва: «Восточная литература», 2001. – С.208-218.

Надійшла: 03.08.2019 Рецензент: д.т.н., доцент Гайдур Г.І.