

УДК 621.37

К.т.н., доцент Маслянюк П.П., студент Якименко П.А.

**Національний технічний університет України
«Київський політехнічний інститут»**

МОДЕЛЬ ВИЯВЛЕННЯ ПРИХОВАНИХ ЗВ'ЯЗКІВ МІЖ СУТНОСТЯМИ СКЛАДНИХ МЕРЕЖ

Abstract

*Pavlo P. Maslyanko, associate professor, PhD; Pavlo A. Yakimenko, student
Definition model of hidden relations between essences in complex networks*

The study of complex networks is a young and active area of scientific research inspired largely by the empirical study of real-world networks such as computer networks and social networks. This paper concerns the task of definition relations in complex networks.

Вступ

Складні мережі трапляються практично у всіх аспектах людської взаємодії. Прикладами складних мереж є [1, 2]:

1. Соціальні мережі.
2. Мережі цитування.
3. Мережі громадського транспорту.
4. Комп'ютерні мережі:
 - World Wide Web;
 - Internet (мережа маршрутизаторів, мережа e-mail листів).

Система – це множина сутностей і відношень між ними. Сутності і відношення можуть проявлятися в математичному або в нематематичному вигляді [3].

Мережа – система зв'язку сутностей. Для передачі інформації можуть бути використані різні фізичні явища. Загальним випадком мережі є складна мережа. Робота присвячена проблемі виявлення прихованих зв'язків між сутностями складних мереж.

Тема дослідження є доволі новою, і фактично єдиним джерелом інформації є база даних наукових препринтів arXiv.org [4].

Постановка задачі

Об'єкт дослідження-складна мережа.

Предмет дослідження - моделювання прихованих зв'язків між сутностями складних мереж.

Мета дослідження - виявлення прихованих зв'язків.

Проаналізувати реальну складну мережу, виділити основні її характеристики, побудувати матрицю прихованості, виділити приховані зв'язки, проаналізувати отримані результати (присутність прихованих зв'язків та основні характеристики у цьому контексті).

Термінологія

Складна мережа (Complex network) - мережа з нетривіальними топологічними властивостями, які не властиві звичайним мережам (випадковий граф Ердоша-Рені та ін.) [1, 2].

Ступінь вузла k – кількість зв'язків, приєднаних до цього вузла.

Матриця суміжності (adjacency matrix) A - квадратна матриця $N \times N$, де її елементи $a_{ij} = 1$, коли вузли i, j з'єднані між собою, та $a_{ij} = 0$ коли не з'єднані [2].

Виявлення прихованих зв'язків

Для виділення основних характеристик обчислимо наступні параметри [2,3]:

- 1) Ступінь вузла k – кількість зв'язків, приєднаних до цього вузла

$$k_i = \sum_{j=1}^N a_{ij}.$$

- 2) Максимальний найкоротший шлях

$M_1 = \max\{l_{ij}\}$, де l_{ij} - довжина найкоротшого шляху між i, j .

- 3) Середній найкоротший шлях

$$\langle l \rangle = \frac{2}{N(N-1)} \sum_{\substack{i=1 \\ i>j}}^N l_{ij}.$$

- 4) Коефіцієнт кластерності вузла-показує, скільки найближчих сусідів заданого вузла є найближчими сусідами один для одного

$$C_m = \frac{2E_m}{k_m(k_m-1)}.$$

- 5) Загальний коефіцієнт кластерності-вказує на ймовірність існування зв'язку між двома випадково взятими найближчими сусідами вузла

$$C = \frac{1}{9} \frac{\sum_{\substack{i=1 \\ i \neq j}}^N (A^3)_{ii}}{\sum_{\substack{i=1 \\ i \neq j}}^N (A^2)_{ij}}.$$

- 6) Посередництво вузла m (betweenness) – величина, що відображає роль вузла в установленні зв'язків і показує, скільки найкоротших шляхів проходить через цей вузол

$$\sigma(m) = \sum_{\substack{i=1 \\ i \neq j}}^N \frac{B(i, m, j)}{B(i, j)}, \text{ де } B(i, m, j) \text{ кількість найкоротших зв'язків між}$$

вузлами i, j таких, що проходять через вузол m ,

$B(i, j)$ - загальна кількість найкоротших зв'язків між вузлами i, j .

- 7) Коефіцієнт когезії (cohesion coefficient)

$$G_{ik} = \frac{\det(M)}{\det(M_{(i+k)(i+k)})}, \text{ де елементи матриці } M \text{ обчислюються з матриці}$$

суміжності: $M_{ik} = -A_{ik}$, $M_{ii} = \sum_{\substack{j=1 \\ j \neq i}}^N |A_{ij}|$, та $M_{(i+k)(i+k)}$ - мінор матриці M , що

обчислюється за правилом: рядок i додається до рядка k і

викреслюється, стовпець i додається до стовпця k і викреслюється.

- 8) Середній коефіцієнт когезії

$$G_{av} = \frac{1}{N(N-1)} \sum_{\substack{i,k=1 \\ i \neq k}}^N G_{ik}.$$

- 9) Матриця когезії G_v - матриця, що складається з коефіцієнтів когезії G_{ik} .

- 10) Матриця прихованості F :

$F = K - G_v$, де G_v - матриця коефіцієнтів когезії,

K - матриця, де занулені безпосередньо з'єднані елементи.

Приховані зв'язки будуть представляти додатні порівняно великі значення матриці F [5].

Приведемо розрахунок деяких характеристик «іграшкової мережі 1»: рис. 1

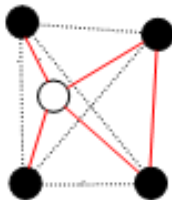


Рис. 1. Граф «іграшкової мережі 1», зв'язки показані жирними лініями

- 1) Максимальна довжина найкоротшого шляху $M_{ij} = 2$.

2) Середня довжина найкоротшого шляху $\langle l \rangle = \frac{3}{2}$.

3) Коефіцієнт кластерності вузла m (відкритий кружечок, $k_m = 4$)

$$C_m = \frac{1}{6}.$$

4) Посередництво вузла m до рівнює $\sigma(m) = 5$.

Для розрахунків коефіцієнтів когезії візьмемо іншу «іграшкову мережу 2»: рис. 2

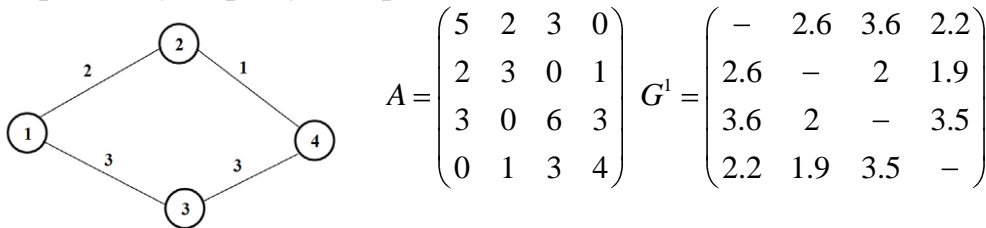


Рис. 2. Граф та модель «іграшкової мережі 2»

Значення середнього коефіцієнту когезії з матриці когезії $G_{av}^1 = 2.2$.

Для демонстрації виявлення прихованих зв'язків розглянемо реальну мережу взаємодії політиків. Виходячи з її розміру, приведемо лише деякі значення, а саме ті, що нас цікавлять.

Номер вузла	Прізвище, що відповідає номеру
3	Басс
4	Тимошенко
9	Яценюк
12	Янукович
27	Кільчицька

Для вибраних персон значення матриці суміжності рівні: $A_{3,3} = 170$, $A_{4,4} = 526$,

$A_{9,9} = 234$, $A_{12,12} = 242$, $A_{27,27} = 20$, звідки випливає, що максимальна кількість зв'язків у Тимошенко. Значення матриці M рівні: $M_{3,4} = 0$, $M_{3,9} = 0$, $M_{3,12} = 0$, $M_{3,27} = 0$, $M_{4,9} = 0$, $M_{4,12} = 0$, $M_{4,27} = 0$, $M_{9,27} = 0$, $M_{12,27} = 0$.

Виявлено, що між вузлами 3 та 9 немає прямого зв'язку, а коефіцієнт когезії цих вузлів $G_{3,9} = 89,4$, що майже в 2 рази більше, ніж загальний коефіцієнт когезії. В цьому випадку можна говорити про прихований зв'язок.

Висновки

В роботі проведено обчислення основних характеристик складних мереж, а також аналіз реальної мережі і виявлення там прихованих зв'язків.

Приховані зв'язки звичайно є не у всіх мережах взаємодії людей, але у великих мережах (наприклад мережа Internet чи мережа взаємодії політиків) будуть присутні обов'язково, і методи, що наведені в цій роботі потрібно буде вдосконалювати чи повністю замінити в залежності від конкретної мережі.

Література

1. *S. N. Dorogovtsev, A. V. Goltsev, J. F. F. Mendes* «Critical phenomena in complex networks» arXiv:0705.0010v6 2007 p. 1-25.
2. *Ю. Головач, О. Олемской, К. фон Фербер, Т. Головач, О. Мриглод, І. Олемской, В. Пальчиков* «Складні мережі» Журнал фізичних досліджень т. 10, № 4 (2006) с. 247–289.
3. *Маслянко П.П.* «Компонентні процеси системного проектування інформаційно – комунікаційних систем». Наукові вісті НТУУ КПІ.- 2008. - №2. - с. 112-121.
4. *Brian Karrer, M. E. J. Newman* «Random graph models for directed acyclic» arXiv:0907.4346v1 2009 p. 1-14.
5. *Снарский А.А., Ландэ Д.В., Женировский М. И.* «Метод выявления неявных связей объектов» Конференция RCDL Петрозаводск, 2009 с. 46-49.