

*Матеріали Міжнародної науково-технічної конференції
«Фундаментальні та прикладні проблеми сучасних технологій», Тернопіль, 2018*

УДК 004.04

Л.В. Омельчук, О.Б. Назаревич

Тернопільський національний технічний університет імені Івана Пулюя

ФОЛКСОНОМІЯ ІНФОРМАЦІЙНИХ РЕСУРСІВ ЯК СКЛАДНА (КОМПЛЕКСНА) БЕЗМАСШТАБНА МЕРЕЖА

L. Omelchuk, O. Nazarevych

FOLKSONOMIA OF INFORMATION RESOURCES AS A COMPLEX (EMPLOYED) EMERGENCY NETWORK

Проаналізувавши явище фолксономії (соціальних закладок) можемо стверджувати, що їх також можна віднести до складних (комплексних) мереж, з одним уточненням – наявності трьох різних типів вершин: користувачі, теги та ресурси з метаданими (веб-сторінки, статті, фото- чи відео-файли). Мережею (network) називається сукупність вузлів, що поєднані зв'язками. Існує багато систем, що мають форму мережі: це інтернет, www, нейронні мережі, мережі метаболізму, харчування, транспортні, розподільчі (наприклад, кровоносні артерії чи поштова служба), соціальні мережі, мережі цитування та соціальні закладки (фолксономія) багато інших. Складною (комплексною) вважається мережа з кількістю більше десятків чи сотні тисяч вершин. Перші статті датовані кінцем 1990-их років. Мета досліджень змінилася від аналізу невеликих графів та властивостей окремих вершин і ребер, до розгляду статистичних властивостей цих графів (мереж).

Зі зміною мети змінилися й методи аналізу. Народження "науки про мережі" відбулося внаслідок розвитку комп'ютерних технологій: www надає порівняно легкий доступ до баз даних з інформацією про різноманітні мережі й водночас потужність комп'ютерів дозволяє докладно їх аналізувати (що було би просто неможливо в іншому випадку, враховуючи розміри досліджуваних мереж). Отож, це був саме той момент, коли наука мала б долучитися як до традиційного емпіричного вивчення мереж, так і до їх моделювання та спроб теоретичного аналізу.

Певний тип мереж, так звані нейронні мережі, цікавив науковців і раніше. Вони й надалі є об'єктом досліджень, проте їхні властивості не настільки визначаються топологією мережі (в принципі, кожен вузол пов'язаний зі всіма іншими), як вагою зв'язків і станом вузлів. Тоді як у поведінці мереж, про які йтиме мова в цьому огляді, топологія відіграє провідну роль. Щоб означити такі об'єкти й коло досліджуваних питань, часто вживається термін складні мережі (complex networks). Відзначимо також споріднені питання, якими цікавляться у фізиці та хімії полімерів складної топології – полімерних мереж.

Виявляється, що найважливіші мережі, які виникають унаслідок людської життєдіяльності, та природні мережі мають специфічну структуру, що характеризується розподілом ступенів вузлів із товстим хвостом (fat-tail distribution) і сильно відрізняється від структури вже дослідженого в математиці класичного випадкового графа. Зазвичай ці мережі не статичні, а такі, що розвиваються, і для розуміння їхньої структури необхідно знати принципи їх еволюції.

Об'єктом, що інтенсивно досліджувався в теорії графів і безпосередньо пов'язаний із складними мережами, є класичний випадковий граф, або випадковий граф Ердоша-Рені. Його ввели і дослідили наприкінці 1950-их Пауль Ердош (Paul Erdos) і Альфред Рені (Alfred Renyi).

Однією з головних характеристик мережі є розподіл ступенів вузлів $P(k)$, що визначається як імовірність того, що вузол i має ступінь $k_i=k$. Під час дослідження виявилось, що мережі, які характеризуються різними $P(k)$, демонструють дуже

різноманітну поведінку, подібно до різних класів універсальності в теорії критичних явищ.

Більшість мереж, що існують у природі, безмасштабні. Чи не найтривалішу історію має вивчення соціальних мереж. Перші спроби виразити соціальні відносини людей у вигляді мережі належать до 30-их років ХХ століття.

Залежно від поставленої задачі, вершинами соціальної мережі можуть виступати окремі особи, групи людей або цілі установи.

Вивчення соціальних мереж було зумовлене, передусім, бажанням дослідити картину взаємодії людей. Окрім того, існує й чисто практичний інтерес до розкриття законів поширення в суспільстві потоків інформації, захворювань, різних уподобань, політичних переконань тощо. При цьому загальні характеристики складних мереж набувають специфічних інтерпретацій при їх застосуванні до конкретної сфери. Для прикладу, ступінь вершини в мережі знайомих показує рівень комунікабельності особи, тобто кількість її соціальних контактів, а середнє значення цього параметра впливає на поширення в суспільстві різноманітних чуток або ж, скажімо, віянь моди. Сукупність методів вивчення механізмів формування соціальних відносин є предметом соціометрії, у межах якої вперше почали конструювати невеликі соціальні мережі: мережа стосунків між школярами (Jacob Moreno, 1934р.); соціальні зв'язки між жінками в одному з міст Південної Америки (Davis, Gardner and Gardner, 1936р.); соціальна мережа робітників заводу в Чикаго (Elton Mayo, 1930р.) та інші.

Найбільшою мережею з доступною інформацією про топологічну структуру є мережа www. Вузли цієї мережі – це документи (веб-сторінки), а спрямовані зв'язки – це гіперлінки, що спрямовують від одного документа до іншого. За станом на січень 2005 р. публічно індексований веб налічував $11.5 \cdot 10^9$ вузлів, причому розмір так званого глибокого чи невидимого вебу (deep web) оцінюється в декілька порядків більшим. Зауважимо, що інформацію про www отримуємо "повзанням" по мережі, а це, своєю чергою, спричиняє недооцінку тих документів, що мають малу цитованість. Відкриття того, що мережа www безмасштабна, сталося наприкінці 1990-их і викликало значне зацікавлення безмасштабними мережами як такими. У результаті аналізу підмножини вебу nd.edu, що складалася на ту пору з 325729 вузлів, отримано in- і out- показники для розподілу ступенів вузлів $Y_{in} = 2.1$, $Y_{out} = 2.45$, а дослідження іншої підмножини, що складалася з $40 \cdot 10^6$ документів, привели до оцінок $Y_{in} = 2.1$, $Y_{out} = 2.38$. Дещо інакше формулювання мережі, коли вузлу ставилось у відповідність ім'я окремого домену, а зв'язок означав, що якийсь із документів цього домену містить покликання на документ іншого домену, привело до оцінки $Y_{in, nom} = 1.94$ [40]. Мережа www виявилася мережею тісного світу з такими оцінками середньої довжини найкоротшого шляху: $\langle l \rangle = 11.2$ (на підставі аналізу 325729 документів) і $\langle l \rangle \sim 19$ (передбачення за допомогою скінчено розмірного скейлінгу для $800 \cdot 10^6$ документів); $\langle l \rangle = 16$ (на підставі аналізу $50 \cdot 10^6$ документів); $\langle l \rangle = 3.1$ (на доменному рівні). Про скорельованість мережі свідчить велике значення коефіцієнта кластерності $C = 0.153127$ (пор. із $C = 0.00023$ для класичного випадкового графа такого ж розміру).

Література:

1. Фолксономия и / или таксономия [Електронний ресурс] / Я.М. Витязев. – режим доступу: <http://blog.vityasev.ru/2007/02/06/folksonomy-and-or-taxonomy/>. – Назва з сторінки Інтернету.
2. Фолксономия [Електронний ресурс] / Wikipedia. – 2015. – режим доступу: <http://ru.wikipedia.org/wiki/Фолксономия>. – Назва з сторінки Інтернету.
3. Фолксономия [Електронний ресурс] / Викизнание. – режим доступу: <http://www.wikiznanie.ru/ru-wz/index.php/Фолксономия>. – Назва з сторінки Інтернету.