

УДК 004.942

В.Пасічник¹, Н.Іванушак²

¹. Національний університет «Львівська політехніка»

². Чернівецький національний університет імені Юрія Федьковича

МОДЕЛЮВАННЯ ПРОЦЕСІВ КЛАСТЕРИЗАЦІЇ КОМП'ЮТЕРНИХ МЕРЕЖ ТА ДОСЛІДЖЕННЯ ЇХ ТОПОЛОГІЧНИХ ВЛАСТИВОСТЕЙ

В даній роботі здійснюється наглядне ймовірнісне моделювання процесу зародження та кластеризації локальної комп'ютерної мережі. Розроблена модель, яка дозволяє отримати зображення мережі для різних початкових умов, динамічно візуалізувати процес її структуризації.

Предметом огляду та дослідження є теорія складних мереж. Форма мережі притаманна багатьом системам, зокрема, - це Інтернет, WWW, нейронні, телекомунікаційні, транспортні, соціальні мережі, мережі цитування тощо.

Складні мережі являються об'єктом як теоретичних, так і емпіричних досліджень [1], в яких топологія розглянутих мереж відіграє провідну роль. Як природні мережі, так і мережі, що виникають внаслідок людської життєдіяльності, зазвичай не являються статичними, а динамічно розвиваються, тому для розуміння їхньої структури необхідно дослідити принципи їх еволюції.

Розглянуті моделі генерації, сформульовані правила структуризації та фактори впливу на динаміку росту складних мереж. На їх основі запропонована імітаційна теоретико-ймовірнісна модель росту локальних комп'ютерних мереж.

В роботах [2,3] означені основні характеристики, які використовуються при дослідженні та моделюванні мереж. «Лінійний розмір» мережі характеризується поняттями середнього $\langle l \rangle$ і максимального l_{\max} найкоротших шляхів. Шляхом між вузлами l_{ij} назвемо найкоротшу відстань між ними. Для зв'язаної мережі з N вузлів середній найкоротший шлях означається як: $\langle l \rangle = \frac{2}{N(N-1)} \sum_{i>j} l_{ij}$, l_{ij} - довжина найкоротшого шляху між вузлами i та j , l_{\max} - найбільше значення з усіх l_{ij} , заданих для цієї мережі.

Головною характеристикою мережі, яка задає розподіл ребер вершини, тобто ступінь вершини, є розподіл ступенів вузлів $P(k)$, що визначає ймовірність того, що вузол i має ступінь $k_i=k$, іншими словами, що випадково вибрана вершина буде мати рівно k ребер. Мережі, які характеризуються різними $P(k)$, демонструють дуже різноманітну поведінку. До найчастіше спостережуваних прикладів розподілу ступенів вузлів відносяться:

a) розподіл Пуассона $P(k) = e^{-\langle k \rangle} \langle k \rangle^k / k!$, (1)

b) експоненційний розподіл $P(k) \sim e^{-k/\langle k \rangle}$, (2)

c) степеневий розподіл $P(k) \sim 1/k^g$, $k \neq 0$, $g > 0$. (3)

В залежності від виду розподілу ступенів вершин (1)-(3) мережі поділяються на три різних типи – класичні випадкові графи, які є варіантом моделі Ердоша-Рені, моделі тісного світу [4] та мережі без масштабування, які мають місце для більшості реальних складних мереж.

В роботі [2] нами були проаналізовані характеристики, динамічні та кореляційні властивості основних типів природних та штучних мереж, таких як соціальні, інформаційні та технологічні, та продемонстровано, що вони являються безмасштабними і підпорядковуються одному і тому ж степеневому закону зростання. Зокрема, до таких мереж відносяться мережі співавторства у різних галузях науки, електронних повідомлень, WWW, мережі цитування, мережі громадського транспорту, Інтернет та ін.

Нами досліджена топологія та розраховані типові характеристики комп'ютерної мережі BW-Star & Fox Net, що знаходиться в місті Чернівці. Для кожного типу вершин, якими є сервери, світчі та користувачі, підраховувалися їхні кількості та ступені, а потім знаходилися кількості $N(k)$ вершин із заданими ступенями k та ймовірність реалізації даного ступеня $P(k)$. Результати досліджень продемонстровані на рис.1, із якого видно що розподіл ступенів

немонотонний і спадає значно повільніше, ніж розподіл Пуассона (1), проте швидше за степеневий розподіл $P(k) = k^{-1.5}$. Це вказує на те, що досліджувана нами мережа займає проміжне місце між класичним випадковим і безмасштабним графами.

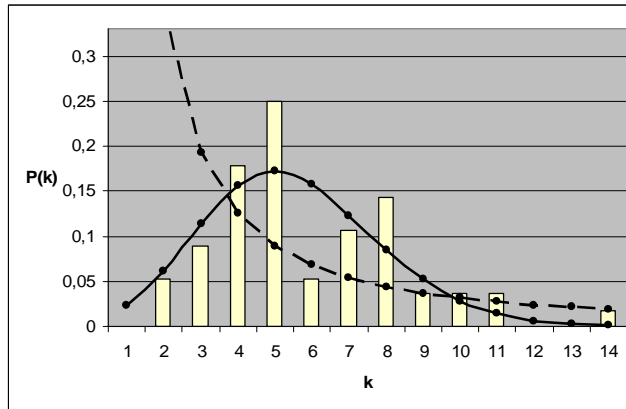


Рис.1. Розподіл ступенів вершин мережі BW-Star & Fox Net (гістограма) в порівнянні з розподілом Пуассона (1) (суцільна лінія) і степеневим законом $P(k) = k^{-1.5}$ (штрихована лінія).

Аналіз розвитку комп'ютерної мережі ґрунтується на встановленні *факторів впливу* на генерацію вузлів та умов утворення і приєднання у мережу нових серверів зі своєю структурою.

Серед факторів впливу на ріст мережі в першу чергу необхідно вирізнити *розмір* або *протяжність* локальної мережі, що визначається відстанню між найбільш віддаленими станціями, при якій за нормальної роботи вузлів чітко розпізнаються колізії, та кількістю об'єднаних в мережу комп'ютерів.

При розростанні мережі зростає число колізій, і різко падає її корисна пропускна здатність та швидкодія передачі сигналу, тому може знадобитися використання дуже дорогого або рідкісного обладнання.

Спостерігається динаміка мережі, своєрідна *кластеризація*, сервери виступають центрами утворених кластерів, відбувається просторове розміщення компонент мережі в чітку ієрархічну структуру.

Враховуючи фактори впливу на ріст та кластеризацію комп'ютерної мережі нами розроблена імітаційна модель, яка дозволяє відтворити фрактальну структуру мережі для різних початкових умов росту, динамічно візуалізувати процес та відслідковувати його в довільний момент часу. Нами розглянута концепція моделювання процесів утворення та росту фрактальних кластерів комп'ютерної мережі за алгоритмом обмеженої дифузії агрегації [5,6] та росту дендритних дерев нейронів.

Перша модельна задача дозволяє на якісному рівні простежити за процесом кластеризації комп'ютерної мережі і дозволяє здійснити оцінки залежності кількості приєднаних у мережу споживачів від радіуса кластера мережі. В якості зародка майбутнього кластера виступає частинка, яку можна ототожнити з *сервером мережі*, до якої в подальшому будуть приєднуватися інші *частинки-споживачі*. Розглянемо стан кожного вузла n . $n = 0$ відповідає стану, коли частка не приєдналася до початкового зародка, а $n = 1$ відповідає приєднанню і утворенню нової структури. Тоді для того, щоб вузол був об'єднаний у структуру необхідно, щоб він знаходився на межі, тобто, мав хоча б одного сусіда з $n = 1$.

Частинки запускаються по одній з довільної випадково вибраної точки на колі заданого радіуса r_0 , побудованого навколо центрального зародка, більшого, ніж радіус кластера r_{max} та слідує по своїх траєкторіях доти, доки не стикнуться із агрегатом.

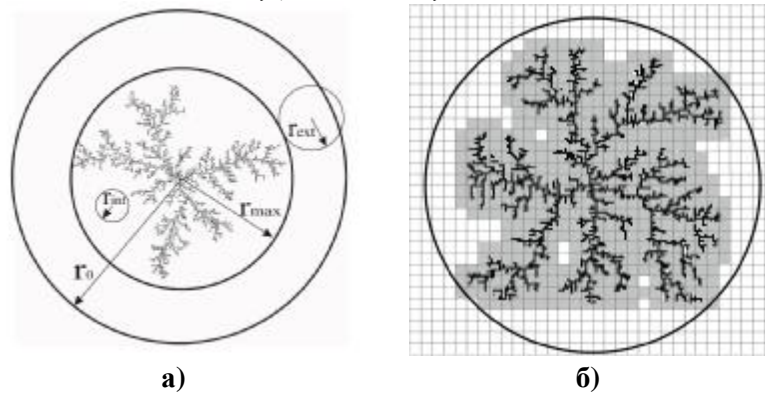


Рис. 2. Схематичне зображення оптимізованих траєкторій.

Якщо в процесі свого руху частинка віддаляється від центру кластера на відстань, що перевищує $3r_0$, то ця частинка вилучається і запускається наступна. У моделі ДООА [6], де частинки випадково блукають, здійснюючи унітарні кроки протягом дискретного проміжку часу, існує стандартний метод, який дозволяє частинкам поза початковим колом здійснювати довгі випадкові кроки довжиною r_{ext} , якщо ці кроки не приведуть частинку всередину початкового кола, як показано на рис. 2 (а). Адекватний вибір – це $r_{ext} = \max(r - r_{max} - d, 1)$, де r - відстань від блукаючої частинки до початкової, і було використано мале допустиме відхилення $d = 5$. Крім того, броунівське блукання у великих порожніх областях у внутрішній частині, яка обмежує кластер ($r < r_{max}$), займає дуже багато комп'ютерного часу, особливо для великих агрегатів. Болл і Брейді [7] запропонували стратегію, яка дозволяє частинкам усередині початкового кола здійснювати великий крок довжини r_{int} , якщо вони не перетинаються у будь-якій частині агрегату, як показано на рис. 2 (а). Аналогічні процедури були використані в інших роботах [5-7].

Моделююча програма працює за наступним алгоритмом: *Створення нової точки. Блукання точки.* (Важлива особливість цієї програми: точка завжди досягає периметра). *Приєднання точки до кластера.* Результатом алгоритму блукання являються координати точки на периметрі кластера. Відбувається збільшення кластера на одну точку у вказаних координатах. Природно, збільшується і периметр кластера.

Проектні рішення:

1. Частинки, що формують кластер, представлені з допомогою класу «точка».
2. Кластер представлений за допомогою лінійного однонапрявленого списку. Елементи списку в проблемній частині містять «точку».
3. Периметр ніяким чином не фіксується і не задається (для економії пам'яті). Для перевірки потрапляння блукаючої частинки на периметр реалізована функція Find Perimeter в класі «кластер».
4. Ймовірність переміщення частинки в одному із восьми напрямків – 1/8. Реалізація випадкового блукання здійснена на основі генератора випадкових чисел з допомогою модуля time.h.
5. На кожному кроці частинка перевіряється на належність до периметра кластера. Якщо частинка знаходиться на периметрі, то вона стає частиною кластера – додається в кінець списку. Пошуковий механізм для визначення того, коли і де блукаюча частинка зіткнеться з агрегатом представляє основний часовий крок у процесі моделювання. Просторові координати частинки, що належить кластеру, зберігаються в одновимірному масиві як послідовності агрегації. Таким чином, перевірка цього масиву здійснюється на кожному кроці моделювання. Для оптимізації пошукового механізму здійснюється перевірка списку у зворотному порядку, в якому частинки були додані в кластер, оскільки для зовнішніх частинок більше шансів для агрегації, ніж для внутрішніх.
6. Для оптимізації часу росту кластера частинка здійснює крок більшої величини при більшій відстані цієї частинки до кластера. Під таке рішення приведені функції DToCluster і SetStep, відповідно, функція розрахунку відстані (найкоротшої) від даного місцезнаходження даної точки до кластера і задання кроку в залежності від цієї відстані. Правильний алгоритм для визначення довжини внутрішніх кроків r_{int} має вирішальне значення для ефективності роботи

програми. Для того, щоб виконати це завдання, ми задаємо квадратну область зі стороною L , зосереджену навколо початкового зародкового зерна, яка обмежує весь агрегат. Ця область повинна бути досить великою для того, щоб гарантувати, що агрегат не перевищує її межі. Тоді, область ділиться на грубозернисту сітку з клітинками розміром $2r_{int} \times 2r_{int}$ як показано на рис. 1(б). Кожна клітинка сітки, асоціюється з елементом $K \times K$ квадратної матриці A , де $K=L/(2r_{int})$, яка має значення 1, якщо клітинка або одна з її найближчих чи наступних за найближчими сусідів містить частку агрегату, або 0 в іншому випадку. Клітки зафарбовані сірим кольором ($A_{ij}=1$) ті, в яких випадкове блукання може перетнути кластер після кроку довжиною r_{int} , так як вони є частиною або суміжні з кластером. Отже, великі кроки, починаючи з сірих клітинок заборонені.

При ДОО окремі частинки після кількох випадкових блукань осідають на кластері, породжуючи випадкові фрактали, які моделюють дендритний ріст комп'ютерної мережі. Для них характерна деревовидна структура з багато чисельними «фьордами» на багатьох розмірних масштабах. Причина утворення подібних структур при ДОО полягає в тому, що блукаюча частинка осідає, як правило, поблизу виступу фракталу, а не в глибині фьорда. Ймовірність проникнення в глибокий фьорд надзвичайно мала.

5	3	4	
3		1	4
3			3
5	3	3	5

Рис. 3

Для обчислення ймовірності прикріплення частинки до кластера відмітимо, що частинка, будучи квадратом, може мати з кластером одну, дві або три спільні вершини. На рис.3 зображений кластер із трьох частин – клітинок з координатами (1;1), (1;2), (2;1), зафарбованих в темний колір. Блукаюча частинка може приєднатися до цього кластера, потрапивши в одну з 12 білих клітинок. Число однокрокових шляхів, якими частинка може потрапити в білу клітинку, записане всередині клітинки. Куткові клітинки мають 1 зв'язок з кластером, серединні – 2 зв'язки, фьорд – 3 зв'язки. Оскільки всього існує 42 однокрокових маршрути, що ведуть в білі клітинки ззовні, то ймовірність p_k того, що частинка потрапить у білу клітинку з k зв'язками:

$$p_1 = 23/42 \approx 0,5476, \quad p_2 = 12/42 \approx 0,4286, \quad p_3 = 1/42 \approx 0,0238.$$

Ми бачимо, що з'єднатися з кластером всередині фьорда у частинки мало шансів. Це означає, що при реальному моделюванні всередині фьорда концентрація частинок менша, так як більшість з них уже приєдналися до кластера. В межах виступів концентрація висока, так як поки що мало частинок з цієї околиці приєдналися до дендриту.

Структура фрактального кластера, змодельованого за даною методикою, зображена на рис.4.



Рис. 4. Структура кластера мережі (для різних значень коефіцієнта дифузії: а – $D=1$, б – $D=0,5$).

Однак запропонована модель не враховує той факт, що комп'ютерна мережа локалізується в структуру з різних часток (світців, споживачів) з різними ступенями приєднання до мережі. Тому далі розглядається модель, яка дозволяє зімітувати ріст комп'ютерної мережі у вигляді дендритів з врахуванням вказаних недоліків.

В ній розглядається мережа як упорядкована множина сегментів, кожен з яких закінчується точкою розгалуження чи кінцем мережі. Вона характеризується низкою числових

характеристик: довжинами сегментів, кутами між сегментами та різними ступенями приєднання вузлів мережі $P(k)$. Початковою точкою мережі вважається сервер, якому приписується z зв'язків, які визначають напрямки $l_0^{(z)}$ зростання мережі. До складу мережі входять два типи часток – світчі та споживачі. Кожному із n світчів випадковим чином приписується різна кількість k зв'язків, яка визначається кількістю портів світча та змінюється дискретно.

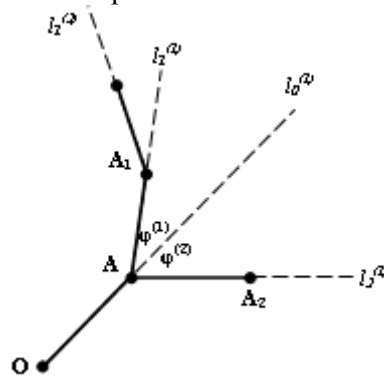


Рис.5. Побудова сегмента комп'ютерної мережі.

Вздовж променя l_0 робиться корок OA певної довжини, точка A з імовірністю $P(k)$ є початком наступного кроку (який відкладається від напрямку попереднього кроку OA під випадковим кутом $j^{(k)}$ відповідно до розподілу кута), і з імовірністю $1 - P(k)$ - не є початком наступного кроку, тобто є кінцем сегмента.

ТАБЛИЦЯ 1

Характеристики комп'ютерних мереж

BW-Star & Fox Net в м. Чернівцях				«Авеню» в м.Сумах [3]		
k	$N(k)$	$P_0(k)$	$P(k)$	$N(k)$	$P_0(k)$	$P(k)$
1	213	0,792	0	631	0,840	0
2	3	0,011	0,053	5	0,007	0,042
3	5	0,019	0,089	18	0,024	0,15
4	10	0,037	0,178	13	0,017	0,108
5	14	0,052	0,25	15	0,020	0,125
6	3	0,011	0,053	13	0,017	0,108
7	6	0,022	0,107	9	0,012	0,075
8	8	0,029	0,143	15	0,020	0,125
9	2	0,007	0,036	3	0,004	0,025
10	2	0,007	0,036	6	0,008	0,05
11	2	0,007	0,036	2	0,003	0,017
12	0	0	0	3	0,004	0,025
13	0	0	0	4	0,005	0,033
14	1	0,004	0,018	6	0,008	0,05
15	0	0	0	4	0,005	0,033
16	1	0,004	0,018	1	0,001	0,008

При програмуванні здійснюється вибірка для реальних комп'ютерних мереж ймовірності приєднання світчів у відповідності до таблиці 1 як розподіл ступенів вершин мереж $P(k)$ ($k = 5, P(k) = 0.25$; $k = 8, P(k) = 0.143$; $k = 16, P(k) = 0.018$).

На основі розроблених алгоритмів реалізована програма, результатом роботи якої являється зображення динаміки росту локальної комп'ютерної мережі, виявлені та проаналізовані особливості роботи запропонованого алгоритму.

Програма допускає коректування форми, розміру, орієнтування у просторі мережі, а також кількості споживачів, які утворюють простір моделювання. Користувачу програми доступні

функції запуску процесу моделювання, його зупинки в довільний момент часу. Програма автоматично генерує зображення та поновлює його після кожного кроку моделювання. Розроблений програмний продукт дозволяє спостерігати за процесом росту мережі на різних стадіях, що дає можливість відслідковувати зміни в структурі не тільки в часі, але і в просторі.

На основі розробленого алгоритму реалізована програма, результатом роботи якої являється зображення динаміки росту локальної комп'ютерної мережі, виявлені та проаналізовані особливості його роботи. Програма допускає коректування форми, розміру, орієнтування у просторі мережі та дозволяє прогнозувати ріст кількості споживачів мережі, які утворюють простір моделювання. Розроблений програмний продукт дозволяє спостерігати за процесом росту мережі на різних стадіях, що дає можливість відслідковувати зміни в структурі не тільки в часі, але і в просторі.

Висновки. В роботі здійснюється імітаційне моделювання процесу росту комп'ютерної мережі у вигляді дендритів. Це пов'язано зі схожістю як у структурному рості, так і в математичному ймовірнісному описі [2] механізмів еволюції складних мереж та росту дендритних кристалів. Розроблена модель, яка дозволяє отримати зображення локальної мережі для різних початкових умов росту, динамічно візуалізувати процес та відслідковувати його в довільний момент часу. Розглянута оригінальна концепція моделювання процесів утворення та росту фрактальних кластерів комп'ютерної мережі за алгоритмом обмеженої дифузії агрегації та росту дендритних дерев нейронів. Отримані результати порівнюються із реальними показниками та характеристиками комп'ютерних мереж.

1. Ю.Головач, О.Олемской, К.фон Фербер, Т.Головач, О.Мриглод, І.Олемской, В.Пальчиков Складні мережі. // Журнал фізичних досліджень. – 2006. – т.10, №4, с. 247-289.
2. Пасічник В.В., Іванущак Н.М. Дослідження та моделювання складних мереж // Східно-Європейський журнал передових технологій. – 2010. – 2/3 (44), с. 43-48.
3. Пасічник В.В., Іванущак Н.М. Структуризація та динамічні властивості складних комп'ютерних мереж. // Східно-Європейський журнал передових технологій. – 2010. – 4/9 (46), с. 16-21.
4. Watts D.J., Strogatz S.H. Collective dynamics of "small-world" networks. // Nature. - 1998. - Vol. 393. pp. 440-442.
5. Liu F., Goldenfeld N. Generic features of late-stage crystal growth // Physical Review A. 1990. V. 42, N 2. P. 895–903.
6. Сандер Л. Континуальна ДОА: случайный фрактальный рост, порождаемый детерминистической моделью / Фракталы в физике. Труды VI международного симпозиума по фракталам в физике (МЦТФ, Триест, Италия, 9 – 12 июля 1985). — М.: Мир, 1988.— С. 336 – 344.
7. R. C. Ball and R. M Brady, J. Phys. A: Math. Gen. 18, L809 (1985). 5. P. Meakin and T. Vicsek, Phys. Rev. A 32, 685 (1986).