Пасічник В. Дослідження еволюційної динаміки та статистичне моделювання комп'ютерних мереж / Пасічник В., Іванущак Н. // Вісник ТНТУ. — 2012. — Том 66. — № 2. — С.151-164. — (приладобудування та інформаційно-вимірювальні технології).

УДК 004.942

В. Пасічник[1], докт. техн. наук; Н. Іванущак[2]

[1]*Національний університет «Львівська політехніка»*
[2]*Чернівецький національний університет імені Юрія Федьковича*

ДОСЛІДЖЕННЯ ЕВОЛЮЦІЙНОЇ ДИНАМІКИ ТА СТАТИСТИЧНЕ МОДЕЛЮВАННЯ КОМП'ЮТЕРНИХ МЕРЕЖ

Резюме. *Розглянуто модель і сформульовано правила структурування складних комп'ютерних мереж із використанням апарата теорії графів. Роботу алгоритму моделювання, адекватність опису моделлю реальної структури проілюстровано шляхом генерації стохастичного графа з використанням статистичних характеристик реальних комп'ютерних мереж. Запропоновані алгоритми моделювання можуть бути використані для розв'язання задачі про стійкість безмасштабних комп'ютерних мереж до спрямованих хакерських атак і розповсюдження комп'ютерних вірусів у них.*

Ключові слова: *комп'ютерні мережі, стохастичний граф, статистичне моделювання.*

V. Pasichnyk, N. Ivanuschak

ANALYSIS OF EVOLUTIONAL DYNAMICS AND STATISTICAL MODELLING OF COMPUTER NETWORKS

The summary. *The object of research carried out in this work is computer systems, the processes of evolutionary dynamics of which are described on the basis of the concept of statistical physics of complex networks. Their topology, structure, hierarchical organization, and the local properties were studied. Computer networks are not static, they evolve in space and that is why for the understanding of the dynamics of their development it is necessary to know the principles of their evolution. From this point of view, the methods were used and the choice of new algorithms for modeling of growth and structure of computer networks were interpreted.*

The model was considered and the rules of structuring of complex computer networks using the apparatus of graph theory were formulated. The algorithm of modeling, the adequacy of description of the real structure by the model is illustrated by the generation of stochastic graph using the statistical characteristics of real computer network, such as "BW-Star & FoxNet" and "DSS-Group" in Chernivtsi and "Avenue" in Sumy.

On the basis of the empirical data and the original method of network generation for the detected law of the network nodes degree distribution modeling in the environment «Processing» were carried out, comparative analysis and systematization of the main characteristics were made. There was analyzed the influence of the statistical characteristics of networks on the structure and properties of stochastic graph models which represent them.

Our estimations, modeling algorithms and the validity of applying mathematical tools allow to make a conclusion on the accuracy and adequacy of the proposed model to the real structures. The proposed algorithms of modeling can be used to solve the problem of the stability of scale-free computer networks against the directed hacker attacks and the distribution of computer viruses.

Key words*: computer networks, stochastic graph, statistical modeling.*

Постановка проблеми. Для сучасних складних систем характерна нерегулярність зв'язків і висока чисельність елементів, яка може досягати десятків і сотень тисяч. Таким системам та їх мережевим моделям, які володіють нетривіальними топологічними властивостями, найбільше відповідає термін «комплексні». Комплексною мережею вважається система, яка (а) складається з великої кількості компонентів, (б) допускає «далекосяжні» зв'язки між компонентами; (в) володіє великомасштабною (у тому числі просторово-часовою) мінливістю. Дана мережа є графом з досить великою кількістю вузлів різної природи, що характеризуються багатовимірним кортежем ознак і динамічно мінливими зв'язками; розподіл ознак вузлів і характеристик зв'язків може бути описаний ймовірнісною моделлю (багатомірним розподілом).

Основною причиною підвищення актуальності розроблень в області теорії і практики комплексних мереж є результати сучасних досліджень реальних комп'ютерних, біологічних і соціальних мереж. Властивості багатьох реальних мереж істотно відрізняються від властивостей класичних випадкових графів з рівноймовірними зв'язками між вузлами, які донедавна розглядалися в якості їх базисного математичного модельного прототипу, і тому побудову їх моделей було запропоновано здійснювати [6, 7] з використанням зв'язних структур і степеневих розподілів.

У теорії комплексних мереж виділяють три основні напрямки:

- дослідження статистичних властивостей, які характеризують поведінку мереж;
- створення моделей мереж;
- прогнозування поведінки при зміні структурних властивостей мереж.

Комплексні мережі використовуються для моделювання об'єктів і систем, дослідження яких іншими способами (за допомогою спостереження або активного експерименту) недоцільні або неможливі.

Комп'ютерні мережі відносяться до мереж, які постійно ростуть і розвиваються. Серед факторів, що впливають на зростання мережі в першу чергу необхідно відзначити розмір або протяжність локальної мережі, яка визначається відстанню між найвіддаленішими станціями, при якій в нормальному режимі роботи вузлів чітко розпізнаються колізії, і кількість об'єднаних у мережу комп'ютерів. Для Інтернет-мереж цей розмір називається діаметром мережі і складає приблизно 1 км відстані, що дозволяє отримати високу швидкість зв'язку та максимально можливий рівень сервісу. При зростанні мережі збільшується кількість колізій, різко падає її корисна пропускна здатність і швидкодія передавання сигналу. Обмеження мережі за довжиною є передумовою вибору структури мережі, розбиття її на окремі частини (сегменти), появи додаткових серверів з новою мережею зв'язків, проблеми генеруються в контексті технологій так званої «останньої милі». Спостерігається динаміка зростання мережі, своєрідна кластеризація, сервери виступають центрами утворених кластерів, відбувається просторове позиціонування компонент мережі у вигляді чітких ієрархічних структур.

Мережа розглядається як множина сегментів, кожен з яких закінчується точкою розгалуження або кінцевої вершиною мережі. Вершинами мережі є сервери, комутатори й кінцеві користувачі, загальну кількість яких позначимо N.

Локальні комп'ютерні мережі є об'єктними прототипами графових структур і тому для їх дослідження застосовують методи теорії графів.

Аналіз останніх досліджень і публікацій. Моделювання мереж із використанням апарата теорії графів є важливим напрямком досліджень дискретної математики [1]. В останні роки зросла зацікавленість дослідників до складних мереж з великою кількістю вузлів, зокрема до комп'ютерних мереж, структура яких нерегулярна, складна і динамічно розвивається в часі [2]. Для таких мереж доводиться генерувати стохастичні графи з величезною кількістю вершин. У загальному вигляді модель комп'ютерної мережі являє собою випадковий граф, закон взаєморозміщення ребер і вершин для якого задається розподілом ймовірностей.

У даний час сформульовано чотири основних підходи до моделювання складних мереж:

- випадкові Пуассонівські графи та узагальнені випадкові графи [3];
- Марковські випадкові графи і модель блукання по «графові графів» із ймовірностями, які пропорційні бажаним властивостям [4];

- модель «тісного світу» Ватса і Строґатса [5] та її узагальнення, еволюційна модель зростання мережі Барабаші й Альберта [6, 7];
- модель Прайса [8].

Перші три підходи передбачають генерацію випадкового графа із заздалегідь відомою кількістю вершин і заданими ймовірнісними властивостями.

Граф Ердоша-Рені є рівноважним ансамблем графів зі сталою кількістю вершин N [3]. Розподіл ступенів вузлів k для цього графа визначається формулою Пуассона $P(k)=e^{-\langle k\rangle}\frac{\langle k\rangle^k}{k!}$. Побудова графа здійснюється генеруванням, коли до N відокремлених вершин послідовно додаються ребра, що з'єднують випадковим чином довільні пари вершин. Початково граф складатиметься із сукупності малих вершин, які в процесі генерування з часом розростаються до гігантського кластера зв'язаних між собою вершин, число яких є скінченною частиною загальної кількості N. При генерації постійно зростає ймовірність зв'язування вершин, яка досягає з часом деякого критичного значення. В результаті процесу, який має характер фазового переходу, граф спонтанно розростається до гігантського кластера вершин, пов'язаних між собою, що нагадує конденсацію краплі води в перенасиченій парі.

Модель Ваттса-Строґаца [5] є комп'ютерною моделлю тісного світу. Її побудова зводиться до наступного: розглядається одновимірний, замкнений у кільце, періодичний ланцюг, який складається із N вершин. Спочатку кожну вершину з'єднують з іншими сусідніми, які знаходяться від неї на відстані, не більшій за k, а потім кожне ребро з певною ймовірністю m перез'єднується з довільною вершиною, що призводить до трансформації регулярного ланцюга у граф тісного світу (рис. 1).

Оскільки в цій моделі кількість ребер є сталою, а ймовірності реалізації графів – різні, то вона зводиться до *канонічного* ансамблю графів і описує реально існуючі мережі, топологія яких не є ані цілком регулярною, ані цілком випадковою.

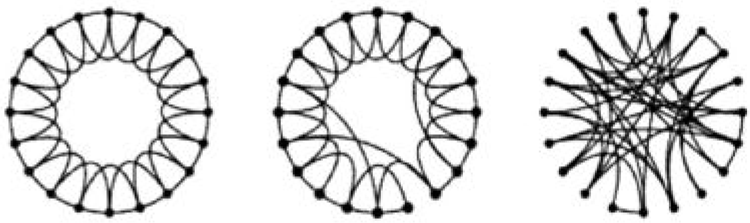

Рисунок 1. Трансформація регулярного ланцюга у граф тісного світу й у випадковий граф
Figure 1. Transformation of the regular chain in the small-world graph and in the random graph

Більшість реальних графів підпорядковуються степеневому закону розподілу $P(k)$. Ці графи побудови мереж описуються моделлю переважного приєднання Барабаші-Альберта [6,7]. Через далекоглядні взаємодії у системи не існує масштабу зміни характерних величин. Ріст і переважне приєднання є основними механізмами побудови безмасштабних мереж.

Нехай вузол *i* має k_i зв'язків і він може бути приєднаним (зв'язаним) до інших вузлів k_i. Ймовірність приєднання нового вузла до вузла i залежить від ступеня k_i вузла i. Величину $W(k_i)=\frac{k_i}{\sum_j k_j}$ називають переважним приєднанням (preferential attachment). Не всі вузли мають однакову кількість зв'язків, тому вони характеризуються функцією розподілу $P(k)$, що визначає ймовірність того, що випадково вибраний вузол має k зв'язків. Для комплексних мереж функція $P(k)$ відрізняється від розподілу Пуассона для випадкових графів. Для переважної більшості комплексних мереж спостерігається степенева залежність $P(k) \sim k^{-\gamma}$.

Метою даної роботи є розв'язання задачі, що виникає при створенні нових методів математичного моделювання для дослідження й обґрунтування підходів, які дозволяють ідентифікувати структуру та параметри моделей комп'ютерних мереж із використанням даних спостережень і теорії комплексних мереж.

Розуміння структури локальної комп'ютерної мережі випливає з дослідження її еволюції в часі, топології й реального розташування. Локальні мережі створюються для оперування у відносно невеликому географічному просторі. Вони дозволяють множинний доступ до високошвидкісного середовища та управління на основі локального адміністрування.

Результати дослідження. Статистичний підхід до опису комплексних мереж. Оскільки випадковий граф містить велику кількість вершин і ребер, то його кількісний опис вимагає використання статистичного підходу [9]. Вершинам графів приписується певне значення енергії E_i, тоді ймовірність P_i того, що вершина має енергію E_i, визначається з умови максимуму ентропії $S = -\sum_i P_i \cdot \ln P_i$. При цьому повинна виконуватися умова нормування $\sum_i P_i = 1$, а внутрішня енергія $U = \sum_i E_i P_i$. Для більшості реальних комплексних мереж кількість вершин і ребер є змінними величинами, тому для опису таких систем використовується великий канонічний ансамбль графів

$$P_a = Z^{-1} \cdot e^{-(E_a - \mu M_a)/T}.$$

P_a задає ймовірність того, що граф зі змінним числом ребер (часток) містить M_a ребер і перебуває в a-мікростані. Статистична сума задається виразом

$$Z = \sum_b e^{-(E_b - \mu M_b)/T},$$

підсумовування здійснюється за всіма мікростанами, тобто за всіма можливими реалізованими графами. Статистика неорієнтованих графів [9] з фіксованою кількістю вершин, які не мають петель, підпорядковується великому канонічному ансамблю, поведінка якого визначається енергією $E = \mu M$, де μ – хімічний потенціал; M – кількість ребер графа. Використовуючи елементи матриці суміжності $A_{i,j} = 0;1$, кількість ребер визначається виразом $M = \sum_{i<j}^{N} A_{ij}$.

Враховуючи його у співвідношенні для енергії, знайдемо статистичну суму

$$Z \equiv \sum_{\{A_{i,j}\}} \exp\left(-\frac{\mu \sum_{i<j} A_{ij}}{T} \right) = \prod_{i<j}^{N} \sum_{A_{ij}=0}^{1} \exp\left(-\frac{\mu A_{ij}}{T} \right) = \left(1 + e^{-\mu/T}\right)^{\mathcal{N}},$$

де ступінь $\mathcal{N} = C_N^2 \equiv \dfrac{N!}{2!(N-2)!}$ визначає кількість пар вершин. У результаті середня кількість ребер визначається за формулою середньостатистичного значення за великим канонічним розподілом

$$\langle M \rangle \equiv \frac{1}{Z} \sum_M M \exp\left(-\frac{\mu M}{T} \right) = -\frac{1}{Z} \frac{\partial Z}{\partial(\mu/T)} = \frac{\partial(F/T)}{\partial(\mu/T)} = \frac{\mathcal{N}}{e^{\mu/T} + 1}.$$

З останнього виразу випливає, що середня кількість ребер $\langle M \rangle$ випадкового графа дорівнює їх загальній кількості $\mathcal{N}$, помноженій на ймовірність реалізації одного зв'язку

$$m \equiv \frac{1}{e^{\mu/T}+1}.$$

Повна ймовірність реалізації графа, з використанням виразу для m, набуває вигляду біноміального розподілу

$$P \equiv \frac{\exp(-E/T)}{Z} = \frac{e^{-\mu M/T}}{\left(1+e^{-\mu/T}\right)^{\mathcal{N}}} = m^M (1-m)^{\mathcal{N}-M}.$$

Як відомо з теорії ймовірностей, у межах нескінченно малого значення $m = \langle M \rangle / \mathcal{N} << 1$ даний розподіл набуває вигляду стандартної форми розподілу Пуассона

$$P = e^{-\langle M \rangle} \frac{\langle M \rangle^M}{M!}.$$

Не зважаючи на малу величину одиничної ймовірності m, в термодинамічній границі $\mathcal{N} = N(N-1)/2 \to \infty$ середня кількість ребер $\langle M \rangle$ виявляється скінченою величиною, яка відіграє роль масштабу розподілу Пуассона. У границі, коли $\langle M \rangle \to \infty$, розподіл за випадковими графами, який визначається кореляціями на довільних масштабах, стає самоподібним, безмасштабним і ймовірність реалізації графа набуває степеневої форми

$$P \sim \langle M \rangle^{-\gamma}.$$

Степеневі розподіли важливі тому, що вони описують самоподібні системи з відсутнім яким-небудь масштабом зміни випадкової величини. Самоподібний розподіл ймовірностей $P(x)$ визначається рівністю $P\left(x/a\right) = a^{\gamma} P(x)$. Використання масштабованої змінної $y \equiv \frac{x}{a}$ і функції $\pi(y) \equiv y^{\gamma} P(y)$ переводить умову самоподібності в однорідний розподіл ймовірностей $P(x) = x^{-\gamma} \pi(y)$, який означає, що зменшення аргументу функції $\pi(y)$ в a разів призводить до факторизації ймовірності випадкової змінної x, піднесеної до степеня $-\gamma$.

Даний статистичний підхід використовується для опрацювання емпіричних даних реальних комп'ютерних мереж, дослідження й моделювання їх росту і структуризації.

Кожен вузол мережі характеризується ступенем k, тобто кількістю зв'язків, які входять у нього. Оскільки в графі, який відображає структуру комп'ютерної мережі, ребра розподіляються випадковим чином, то більшість вершин мають приблизно однаковий ступінь, який близький до його середнього значення $\langle k \rangle$.

Якщо позначити через N загальну кількість серверів, комутаторів і кінцевих користувачів комп'ютерної мережі, а кількість вершин із заданим ступенем приєднання k – через $N(k)$, то ймовірність реалізації цього ступеня k визначається діленням $N(k)$ на загальну кількість вершин $P(k) = N(k)/N$. Останні емпіричні результати свідчать, що для більшості мереж, включаючи Всесвітню павутину WWW та Інтернет, розподіл ступенів вершин є степеневим $P(k) \approx k^{-\gamma}$.

Ймовірнісна модель комп'ютерної мережі. Комп'ютерна мережа зображується у вигляді графа G, який визначається як сукупність (V, E) кінцевої множини вершин V, $\dim(V) = N$, і множини ребер E, яка складається із невпорядкованих пар (u, v), де $u, v \in V$ і $u \neq v$. Кожна вершина характеризується своїм ступенем, тобто числом інцидентних їй ребер. Упорядкований список ступенів вершин називають ступеневою послідовністю.

Інтегральною характеристикою комп'ютерної мережі є закон розподілу ступенів p_k, який задає ймовірність того, що випадково вибрана вершина має ступінь k. Ступеневу послідовність для неорієнтованого графа зручно записати у формі

$$d = \left(k_1^{n_1}, k_2^{n_2}, ..., k_s^{n_s}\right),$$

де числа k_i є ступенями вершин, а показник n_i визначає кількість повторів числа k_i у послідовності. Наприклад, $\left(3^1, 2^2, 1^4\right) = \left(3,2,2,1,1,1,1\right)$. Такий запис дозволяє пов'язати дискретний розподіл ступенів вершин p_k зі ступеневою послідовністю d у формі

$$p_k \overset{def}{=} P\left[x = k_i\right] = n_i / N .$$

У загальному випадку мається на увазі, що ступенева послідовність є монотонно незростаючою, однак у випадку генерації комп'ютерних мереж дана вимога не є обов'язковою.

У моделі випадкових графів [3] ребро, яке інцидентне довільним двом вершинам, присутнє або відсутнє з рівною ймовірністю, а тому розподіл p_k буде біноміальним або (у границі за N) Пуассонівським. Однак більшість реальних мереж має структуру, відмінну від структури випадкових графів, що позначається на характері розподілу ступенів вершин. Зокрема, у багатьох реальних мережах емпіричний розподіл ступенів вершин інтерпретується в термінах ступеневого розподілу $p_k = k^{-\gamma} \big/ \zeta(\gamma)$, де ζ – функція Рімана відіграє роль нормуючої константи. Цей розподіл характеризується єдиним параметром γ, який визначає швидкість спадання «хвоста» розподілу.

Для здійснення процесу моделювання комп'ютерної Інтернет-мережі використовувались характеристики конкретних мереж, а саме, «BW-Star & FoxNet», «DSS-Group» в Чернівцях та «Авеню» в Сумах, наведено в таблиці 1.

Таблиця 1. Характеристики комп'ютерних мереж

	«BW Star & Fox Net» в Чернівцях						«Авеню» в Сумах [9]		«DSS-Group» в Чернівцях	
	2005		2008		2011					
	N=272		*N=588*		*N=915*		*N=747*		*N=2023*	
k	n_k	p_k	n_k	p_k	n_k	p_k	n_k	p_k	n_k	p_k
1	207	0.761	401	0.682	614	0.671	631	0.840	1242	0.613
2	13	0.048	37	0.063	59	0.064	5	0.007	224	0.110
3	5	0.018	48	0.082	79	0.086	18	0.024	220	0.110
4	10	0.036	28	0.047	51	0.056	13	0.017	79	0.040
5	14	0.051	19	0.032	22	0.024	15	0.020	68	0.033
6	3	0.011	14	0.024	33	0.036	13	0.017	35	0.017
7	6	0.022	6	0.010	9	0.010	9	0.012	46	0.022
8	8	0.029	8	0.014	8	0.009	15	0.020	26	0.012
9	2	0.007	8	0.014	10	0.011	3	0.004	13	0.006
10	2	0.007	6	0.010	6	0.006	6	0.008	21	0.010
11	2	0.007	6	0.010	8	0.009	2	0.003	10	0.005
12	0	0	2	0.003	7	0.008	3	0.004	14	0.007
13	0	0	2	0.003	3	0.003	4	0.005	8	0.004
14	1	0.004	1	0.001	3	0.003	6	0.008	9	0.004
15	0	0	2	0.003	3	0.003	4	0.005	4	0.002

Ступінь вузла k задає кількість ребер інцидентних конкретній вершині, а n_k – кількість вершин у графі із заданим k. За цими даними побудований розподіл ступенів вершин подано у вигляді гістограми на рис. 1. На рис. 2 здійснена апроксимація «хвостів» розподілів ступенів вершин досліджуваних мереж і встановлені для них значення параметра степеневого розподілу $p_k = k^{-2.2}$ для мережі «BW-Star & Fox Net» та $p_k = k^{-1.5}$ для мережі «Авеню».

Для здійснення процесу моделювання проводиться вибірка ймовірностей приєднання вузлів як розподілів ступенів вершин мереж p_k для реальних комп'ютерних мереж відповідно до таблиці 1.

У процесі досліджень ми мали можливість простежити за розвитком та структуризацією комп'ютерної мережі «BW-Star & Fox Net» у часі. Характеристики цієї мережі в різні часові проміжки наведено в таблиці 1. Якщо на початку становлення мережа займала проміжне місце між масштабною з Пуассонівським розподілом ступенів вершин та безмасштабною мережами, то з часом відбувається зростання та структуризація системи, розподіл ступенів вершин для неї вже інтерпретується в термінах степеневого розподілу $p_k = k^{-\gamma}$, причому значення показника $\gamma = 2{,}4$ практично залишається незмінним за останні роки в період з 2008 до 2011 року. Це вказує на те, що топологічні властивості мережі вже перебувають у стійкому стаціонарному стані.

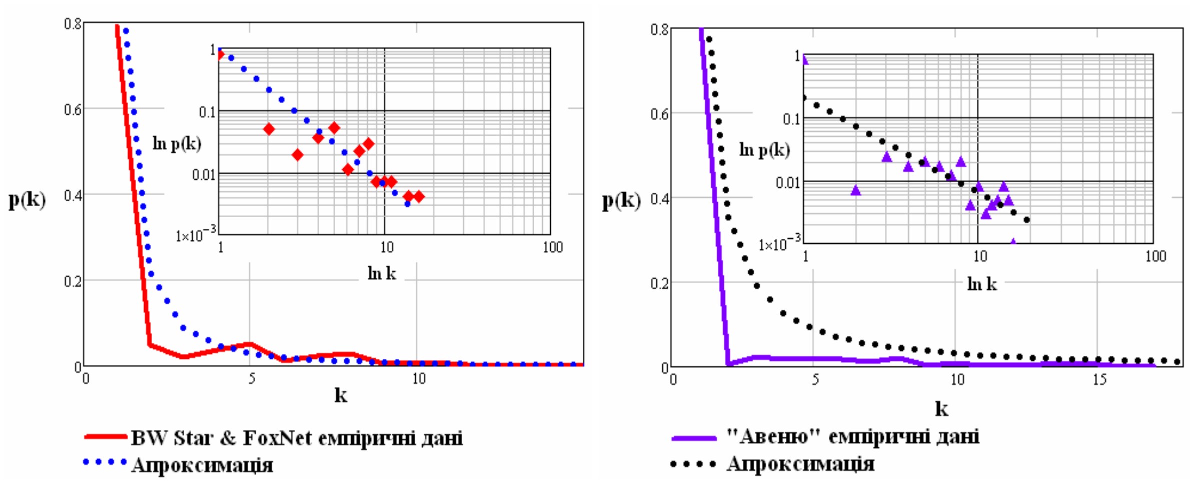

Рисунок 2. Апроксимація «хвостів» розподілів ступенів вершин досліджуваних мереж
Figure 2. Approximation of "tails" of degree distributions of vertices in investigated networks

Отже, якщо мережа достатньо структуризована, то в процесі її розвитку змінюється кількість вершин n_k із заданим ступенем вузла k, рівно як і загальна кількість N користувачів мережі. Однак ймовірності приєднання цих вершин p_k залишаються практично незмінними, забезпечуючи тим самим степеневий розподіл ступенів вершин $p_k = k^{-\gamma}$ з незмінним показником степеня γ.

Спосіб генерації комп'ютерної мережі для заданого закону розподілу ступенів вершин.

Параметри узагальненої конфігураційної моделі:

N – число вершин у мережі;

s – число класів вершин;

$i \in \{1,2,3,...,s\}$ – позначає конкретний клас вершин;

n_i – число вершин *i*-го класу;

k_i – ступінь вершини *i*.

Так як розподіл ступенів вершин p_k заданий, то зведемо обчислювальну процедуру до таких операцій:

• сформуємо ступеневу послідовність d, вибираючи s чисел n_i згідно із заданим розподілом p_k, де $i = \overline{1, s}$;

• кожній вершині i графа присвоїмо k_i «заготовок» (кінців) для майбутніх ребер;

• зі ступеневої послідовності випадково отримуються пари «заготовок». Вони з'єднуються ребром у випадку, якщо нове ребро не призведе до утворення ребер-циклів (петель) або мультиребер. Якщо ребро згенероване, то відповідні індекси зі ступеневої послідовності видаляються;

• попередній крок повторюється до тих пір, поки ступенева послідовність не стане порожньою;

• укладка графа здійснюється шляхом розміщення вершин із найбільшими ступенями приєднання в центрі графа, а вершини з меншими ступенями радіально розташовуються від центру до периферії в порядку зменшення їх ступенів;

• зв'язки між вузлами заповнюються послідовно, починаючи з вершин з найбільшою кількістю ребер.

Остання умова забезпечує об'єднання всіх вузлів у єдину структуру стохастичного графа, що відображає факт обов'язкового приєднання всіх користувачів у реальну локальну комп'ютерну мережу.

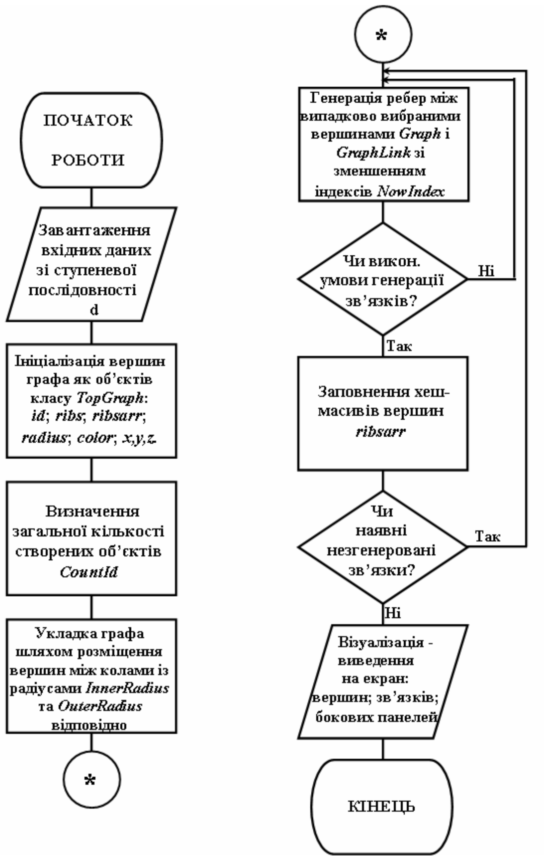

Рисунок 3 Блок-схема програмної реалізації алгоритму моделювання
Figure 3. Block diagram of program realization the algorithm of modeling

На основі розподілу p_k довільний граф може бути побудований $\prod_i k_i!$ різними способами, так як «заготовки» для майбутніх ребер нерозрізнимі. Таким чином, цей процес з рівною ймовірністю генерує довільну можливу конфігурацію мережі із заданим розподілом ступенів вершин p_k.

Перевагою даного алгоритму є його універсальність, з його допомогою можливо побудувати мережу з довільним розподілом ступенів вершин.

Наведемо приклад побудови графа за сформованою ступеневою послідовністю у вигляді

$$d = \left(3^2, 2^3, 1^4\right).$$

Перший крок – нумеруються усі вершини із заданої послідовності. У центрі графа розміщуються вершини з найбільшим ступенем вершин $k = 3$, навколо яких послідовно розташовуються вершини з нижчими ступенями приєднання вузлів. Наступним кроком є генерація ребер E, тобто процес випадкового з'єднання пар заготовок вершин. При цьому відслідковуються утворення нових ребер, щоб у графі не з'явилися мультиребра. Це здійснюється завдяки тому, що при генерації ребра індекси, які йому відповідають, зі ступеневої послідовності видаляються і формуються правила заборони, тобто для кожної вершини зберігаються номери усіх суміжних з нею вершин для того, щоб не дозволити повторного з'єднання вузлів, між якими уже згенероване ребро.

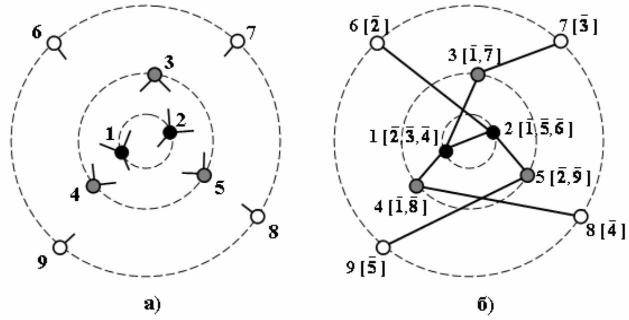

Рисунок 4. Побудова графа за заданою ступеневою послідовністю

Figure 4. Construction of the graph for a given graduated sequence

На рис. 4а кожній вершині графа присвоєно k_i (i = 1, 2, 3) «заготовок» для майбутніх ребер, на рис. 4б зображено граф, згенерований одним зі способів. Числа в квадратних дужках демонструють заборону з'єднання з вузлом відповідного номера після утворення ребра. Граф, який відповідає ступеневій послідовності (2), може бути побудований $1! \cdot 2! \cdot 3! = 12$ різними способами.

Даний алгоритм реалізований у Processing, який є загальнодоступною мовою програмування й середовищем для високоякісної візуалізації зображень, анімацій та їх взаємодій. Являє собою предметно-орієнтовану мову програмування, засновану на java з простим Сі-подібним синтаксисом. Створений як елемент для основ програмування у контексті візуалізації, у розпорядженні інструменти для побудови графічних, 3D-об'єктів, робота зі світлом, текстом, інструментами трансформації.

Комп'ютерний експеримент. Результатом програмної реалізації запропонованого алгоритму є власне комп'ютерна мережа, зображена у вигляді стохастичного графа з відомим числом вершин і заданим розподілом ймовірностей їх приєднання.

Робота алгоритму моделювання, адекватність описання моделлю реальної структури проілюстрована шляхом генерації графа з використанням характеристик реальних комп'ютерних мереж BW-Star & Fox Net та DSS-Group у Чернівцях. За вибіркою визначено розподіли таких числових характеристик для реальних мереж:

1) кількість вершин у мережі n_k з різними ступенями їх приєднання;
2) впорядкований список ступенів вершин у вигляді ступеневої послідовності $d = \left(k_1^{n_1}, k_2^{n_2}, ..., k_s^{n_s}\right)$ для моделювання стохастичного графа;
3) відповідні ймовірності $p_1, p_2, p_3, ..., p_s$ приєднання вершин з різними ступенями k_i ($i = \overline{1, s}$) у мережу.

Вибірка здійснювалася за емпіричними розподілами ступенів вершин, які інтерпретуються в термінах степеневого розподілу $p_k = k^{-\gamma}$. На її основі здійснено процес моделювання мережі з подальшою можливістю порівняння результатів моделювання з характеристиками досліджуваних мереж, наведеними у табл. 1, та оцінювання адекватності описання моделлю реальної структури.

Провівши апроксимацію «хвостів» розподілів ступенів вершин, що проілюстровано на рис. 2, та визначивши тим самим показники γ для різних локальних комп'ютерних мереж, зокрема $\gamma = 2{,}4$ для мережі «BW-Star & Fox Net», $\gamma = 1{,}5$ для мережі «Авеню» та $\gamma = 2{,}1$ для мережі «DSS-Group», здійснено моделювання цих мереж за відомим показниковим $p_k = k^{-\gamma}$ розподілом ймовірностей приєднання користувачів у мережу з відповідними γ. Як ілюстрація, на рис. 5 наведено приклади візуалізації стохастичних графів, які відображають властивості досліджуваних комп'ютерних мереж. Для здійснення процесу динамічної візуалізації використовувався оригінальний алгоритм укладання графа, який, на нашу думку, дає найбільш інформативне відображення структури та властивостей комп'ютерних мереж.

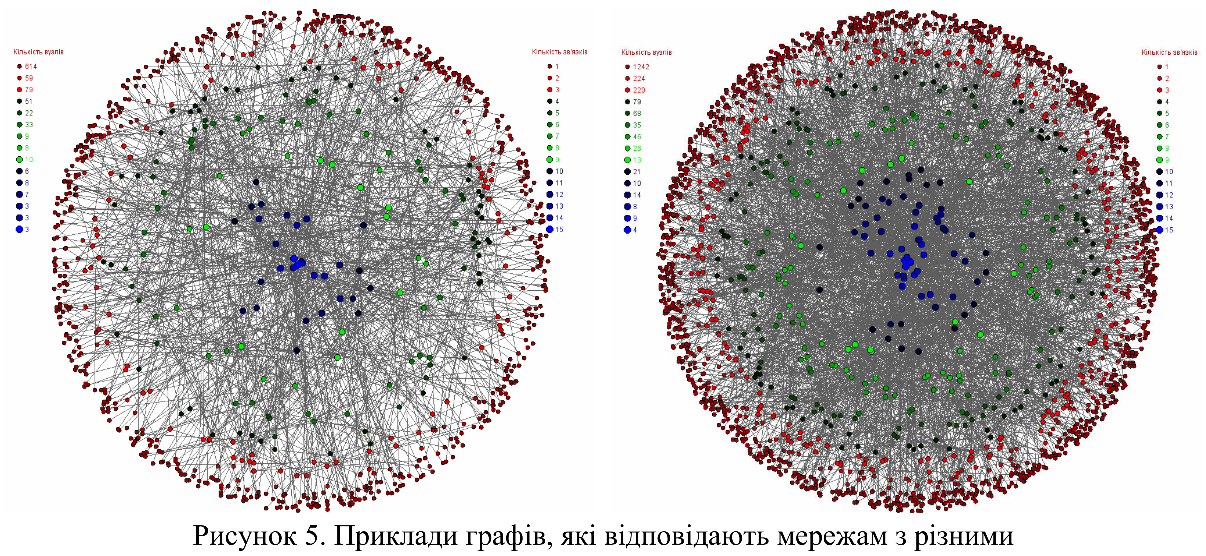

Рисунок 5. Приклади графів, які відповідають мережам з різними законами залежності $p_k = k^{-\gamma}$:

(а) $\gamma = 2{,}4$, $N = 915$; (б) $\gamma = 2{,}1$, $N = 2023$

Figure 5. Examples of graphs that correspond to networks with different laws depending $p_k = k^{-\gamma}$:

(а) $\gamma = 2{,}4$, $N = 915$; (б) $\gamma = 2{,}1$, $N = 2023$

На рис. 5 вершини з різними ступенями приєднання k зображені різними кольорами, їх кількість у згенерованій мережі винесено на панель зліва, кількість зв'язків, які відповідають кожній вершині, відображені на панелі справа.

З рис. 5 бачимо, що для малих значень параметра розподілу γ мережа кластеризується в один більший зв'язаний кластер, аніж у випадку з більшими значеннями γ. Завдяки тому, що основний внесок у мережу роблять користувачі, для яких ступінь приєднання у мережу $k=1$, то середній ступінь мережі

$$\langle k \rangle = \sum_k k \cdot p_k \ ,$$

знайдений у такий спосіб, є порівняно малою величиною. Для мережі «BW-Star & Fox Net» його значення $\langle k \rangle = 1{,}997$. Слід відзначити, що для переважної більшості комп’ютерних Інтернет-мереж середній ступінь $\langle k \rangle$ в міру тих же причин прийматиме малі значення. Щодо параметра γ показника степеня степеневого розподілу, то його значення може бути різним. Більшим – у менш розгалужених системах з порівняно малою кількістю серверів, світчів з багатьма зв’язками k і навпаки – великою кількістю користувачів з $k=1$. Меншим – у більш структуризованих мережах, у структурі яких є достатня кількість вершин з великими ступенями k (рис. 5б), таких, як мережа DSS-Group у Чернівцях.

Для обґрунтування результатів комп’ютерного експерименту обчислені середні значення коефіцієнтів кластерності систем

$$\langle C \rangle = \frac{c_k n_k}{N} \ ,$$

де c_k – локальна величина коефіцієнта кластерності. Для мережі «BW-Star & Fox Net» $\langle C \rangle = 0{,}032$, для мережі «Авеню» $\langle C \rangle = 0{,}081$. Мале значення $\langle C \rangle$ вказує на низьку кореляцію у локальних комп’ютерних мережах. Згідно з результатами комп’ютерного експерименту середній ступінь вузла $\langle k \rangle$ і коефіцієнт кластерності $\langle C \rangle$ мають тенденцію до повільного збільшення при розростанні мережі.

Запропонований алгоритм та його програмна реалізація дозволяють генерацію комплексних мереж з природною кількістю зв’язків у них $N \sim 10^3 \div 10^5$.

Результати дослідження. Необхідно знати й уміти моделювати не тільки структуру зв’язків у даний момент часу, але й динаміку мережі з конкретним розподілом зв’язків за достатньо великий проміжок часу. Запропонований у роботі алгоритм дозволяє здійснити прогнозування розвитку мережі. Як приклад, відслідковуючи динаміку становлення мережі «BW-Star & Fox Net» за останні роки, наведену в табл. 1, здійснивши процес моделювання за визначеним для неї інтегральним законом розподілу ймовірностей $p_k = k^{-2,4}$ та щорічним приростом $\Delta N \approx 110$ зв’язків цієї мережі, можна обчислити прогнозовану кількість користувачів, серверів та комутаторів у ній у 2014 р (табл. 2).

Таблиця 2. Прогнозована кількість зв’язків у мережі

$N = 1245$	k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	n_k	751	271	89	45	26	17	12	8	6	5	4	3	3	3	2

Результати, отримані при вивченні моделі, можна переносити на реальну структуру, якщо модель її адекватно описує. Питання про адекватність, точність та достовірність змодельованої системи і досліджуваних реальних комплексних мереж вивчалося шляхом зіставлення, порівняння та оцінювання їх числових характеристик (кількості вузлів з різними ступенями k, ймовірності об’єднання у мережу). За міру розкиду даних вибиралася дисперсія або середній квадрат відхилення (σ^2), який характеризує відхилення випадкових значень від середньої величини в даній вибірці. У роботі питання про адекватність моделі встановлюється через зіставлення оцінок усередненого апроксимованого та реального розподілів ступенів вузлів p_k згідно з формулою

$$\sigma^2 = \frac{1}{n}\sum_{k=1}^{n}(p_k - \overline{p})^2 .$$

Розкид даних для мережі «BW-Star & Fox Net» у різні часові проміжки, наведених у табл. 1, порівняно з усередненими апроксимованими оцінюється значеннями $\sigma^2 = 0.0057$ для даних 2005 року, $\sigma^2 = 0.0013$ – для даних 2008 року та $\sigma^2 = 0.0012$ – для даних 2011 року, що у відсотковому відношенні складає $7.5 \div 3.5$% мінливості варіаційного ряду.

На основі проведених числових розрахунків можна дійти висновку про адекватність описання моделлю реальних мереж.

Такі моделі використовуються, зокрема, для опису розповсюдження епідемій (наприклад, таких, як грип або ВІЛ) у соціальних мережах [10]. Алгоритми моделювання, запропоновані в даній роботі, мають стати засобом для вироблення підходів до діагностики процесів розповсюдження комп’ютерних вірусів у комп’ютерних мережах і дослідження вразливості та стійкості останніх до спрямованих атак.

Аналіз реальних безмасштабних мереж WWW та Інтернету [11, 12], метаболізму [13], мережі харчування [14] демонструє неабияку їх стійкість до вилучення вузлів: тобто ці мережі виявляють несподіваний ступінь стійкості при випадкових ураженнях. З іншого боку, при спланованих сценаріях нанесення шкоди або вірусних атаках, мережа стає надзвичайно вразливою.

Висновки. Завдяки дослідженням вивчено підхід до моделювання динаміки розвитку та становлення комп’ютерних мереж із використанням апарату складних мереж. У рамках запропонованої схеми розроблено програмне забезпечення у середовищі Processing для моделювання комплексних комп’ютерних мереж.

В ході роботи проаналізовано вплив статистичних характеристик мереж на структуру та властивості модельних стохастичних графів, які їх зображають. Сформульований у роботі підхід до моделювання дає можливість згенерувати випадкові графи з відомим заздалегідь числом вершин і заданими ймовірнісними властивостями.

Проведені оцінювання, використані алгоритми моделювання та обґрунтованість застосування математичного апарату дозволяють зробити висновок про точність та адекватність запропонованої моделі до реальних структур.

Запропоновані в роботі алгоритми моделювання можуть бути використані для розв’язання задачі про стійкість безмасштабних комп’ютерних мереж до спрямованих хакерських атак і розповсюдження комп’ютерних вірусів у них.

Conclusions. The approach to the modeling of the development dynamics and formation of the computer networks using the apparatus of complex networks has been studied due to our investigations. Under the proposed scheme software in Processing environment for the modeling of complex computer networks has been developed. The approach of modeling presented in paper makes it possible to generate random graphs with known pre-specified number of vertices and probabilistic properties.

Our estimations, modeling algorithms and the validity of applying mathematical tools allow to make a conclusion on the accuracy and adequacy of the proposed model to real structures.

The suggested simulation algorithms can be used to solve the problem of stability of scale-free computer networks to targeted hacker attacks and the spread of computer viruses in them.

Список використаної літератури

1. Нікольський, Ю. Дискретна математика [Текст] / Ю.В. Нікольський, В.В. Пасічник, Ю.М. Щербина. – Львів: «Магнолія-2006». – 2009. – 432 с.

2. Newman, M.E.J. The Structure and Function of Complex Networks [Text] / M.E.J. Newman // SIAM Review. – 2003. – Vol. 45. – N. 2. – P. 167–256.

3. Erdős, P. On the evolution of random graphs [Text] / P. Erdős, A. Renyi // Publications of the Mathematical Institute of the Hungarian Academy of Sciences. – 1960. – Vol. 5. – P. 17–61.

4. Frank, O. Markov graphs [Text] / O. Frank, D. Strauss D // Journal of the American Statistical Association. – 1986. – Vol. 81. – P. 832–842.

5. Watts, D.J. Collective dynamics of "small-world" networks [Text] / D.J. Watts, S.H. Strogatz // Nature. – 1998. – Vol. 393. – P. 440–442.

6. Albert, R. Statistical mechanics of complex networks [Text] / R. Albert, A.-L. Barabasi // Reviews of Modern Physics. – 2002. – P. 47–97.

7. Barabasi, A-L. Emergence of scaling in random networks [Text] / A.-L. Barabasi, R. Albert // Science. – 1999. – Vol. 286. – P. 509–512.

8. Price, D.J. de S. A general theory of bibliometric and other cumulative advantage processes [Text] / D.J. de S. Price // Journal of the American Society for Information Science. – 1976. – Vol. 27. – P. 292–306.

9. Головач, Ю. Складні мережі [Текст] / Ю. Головач, О. Олємской, К.фон Фербер, Т. Головач, О. Мриглод, І. Олємской, В. Пальчиков // Журнал фізичних досліджень. – 2006. – Т.10. – № 4. – С. 247 – 289.

10. Sloot, P.M.A. Stochastic simulation of HIV population dynamics through complex network modeling [Text] / Sloot P.M.A., Ivanov S.V., Boukhanovsky A.V., D.A.M.C. van de Vijver c and C.A.B. Boucherc // International Journal of Computer Mathematics. – 2008. – Vol. 85. – N. 8. – P.1175–1187.

11. Albert, R. Error and attack tolerance of complex networks [Text] / R. Albert, H. Jeong, A.-L. Barabasi // Nature (London). – 2000. – Vol. 406. – P. 378–381.

12. Tu, Y. How robust is the Internet? [Text] / Y. Tu // Nature (London). – 2000. – Vol. 406. – P. 353–354.

13. Jeong, H. The large-scale organization of metabolic networks [Text] / H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, A.– L. Barabasi // Nature (London). – 2000. – Vol. 407. – P. 651–654.

14. Sole, R. V. Complexity and fragility in ecological networks [Text] / R.V. Sole, J.M. Montoya // Proc. R. Soc. Lond. – 2001. – B 268. – P. 2039–2045.

Отримано 26.03.2012