УДК 004.942

В.В. ПАСІЧНИК, Н.М. ІВАНУЩАК

МОДЕЛЮВАННЯ І ВІЗУАЛІЗАЦІЯ ЗАСОБАМИ PROCESSING КОМП'ЮТЕРНИХ МЕРЕЖ

Анотація. *У роботі представлені дослідження комп'ютерних інтернет-мереж на основі концепції статистичної фізики складних мереж, теоретичні обґрунтування та симуляції з використанням математики як інструменту та мови аналізу, розглянуто методику та здійснено моделювання росту й структуризації локальних мереж, результати якого узгоджуються з емпіричними даними.*
Ключові слова: *комп'ютерні мережі, стохастичний граф, статистичне моделювання.*

Аннотация. *В работе представлены исследования компьютерных интернет-сетей с использованием концепции статистической физики сложных сетей, теоретические обоснования и симуляции с использованием математики как инструмента и языка анализа, рассмотрена методика и произведено моделирование роста и структуризации локальных сетей, результаты которого согласуются с эмпирическими данными.*
Ключевые слова: *компьютерные сети, стохастический граф, статистическое моделирование.*

Abstract. *The paper presents the investigation of computer internet-networks, based on the concept of statistical physics of complex networks, theoretical studies and simulations using mathematics as a tool and analysis language; the technique was considered and modeling of the growth and structuring of local networks was performed. The results of this growth are coordinated with the empirical data.*
Keywords: *computer networks, stochastic graph, statistical modeling.*

1. Вступ

Метою роботи є рішення задачі, яка має суттєве значення для створення нових методів математичного моделювання, а саме: дослідження і обґрунтування підходів, які дозволяють ідентифікувати структуру і параметри моделей локальних комп'ютерних мереж на основі даних спостережень.

Основною причиною актуальності теорії складних мереж є результати сучасних робіт з опису реальних комп'ютерних, біологічних і соціальних мереж. Властивості багатьох реальних мереж суттєво відрізняються від властивостей класичних випадкових графів з рівноймовірними зв'язками між вузлами, і тому вони будуються на основі зв'язаних структур та степеневих розподілів.

У теорії складних мереж виділяють три основних напрями:

– дослідження статистичних властивостей, які характеризують поведінку мереж;

– створення моделей мереж;

– прогнозування поведінки мереж при зміні їх структурних властивостей.

Складні мережі застосовуються для моделювання об'єктів і систем, для яких інші способи дослідження (за допомогою спостереження і активного експерименту) є недоцільними або неможливими.

2. Моделювання комп'ютерної мережі з використанням апарата теорії графів

Моделювання мереж з використанням апарата теорії графів є важливим напрямом 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 імовірностей.

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

– випадкові пуасонівські графи та узагальнені випадкові графи [3];

– марковські випадкові графи і модель блукання по «графові графів» з імовірностями, які пропорційні бажаним властивостям [4];

– модель «тісного світу» Ватса і Строгатса [5] і її узагальнення, еволюційна модель росту мережі Барабаші і Альберта [6];

– модель Прайса [7].

Перші три підходи передбачають генерацію випадкового графа із заздалегідь відомим числом вершин і заданими ймовірнісними властивостями.

3. Імовірнісна 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. Впорядкований список ступенів вершин називається 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 подати 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. Так, наприклад, $\left(3^1, 2^2, 1^4\right) = (3,2,2,1,1,1,1)$. Такий запис дозволяє пов'язати дискретний розподіл ступенів вершин p_k зі ступеневою послідовністю d у формі

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

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) пуасонівським. 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 $p_k = k^{-\gamma} / \zeta(\gamma)$, 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 м. Чернівцях та «Авеню» в м. Сумах, подані в табл. 1. 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, поданий у вигляді гістограми на рис. 1. На рис. 2 здійснена апроксимація «хвостів» розподілів ступенів вершин досліджуваних мереж та 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».

Таблиця 1. Характеристики комп'ютерних мереж

	«BW Star & Fox Net » в м. Чернівцях						«Авеню» в м. Сумах [8]		«DSS-Group» в м. 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

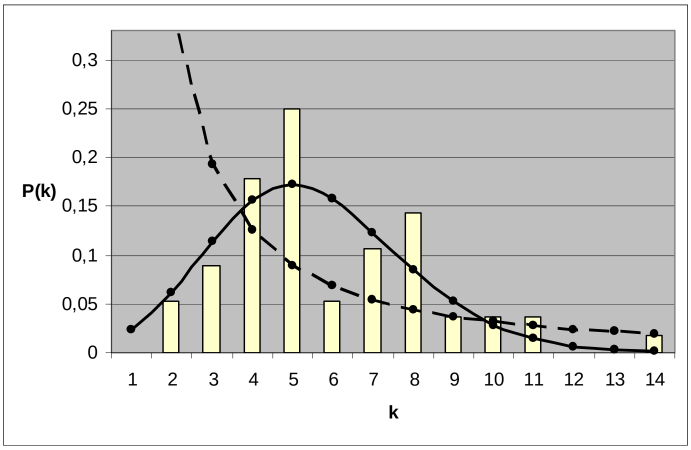

Рис. 1. Розподіл ступенів вершин мережі «BW-Star & Fox Net» (гістограма) в порівнянні з розподілом Пуассона (суцільна лінія) і степеневим законом $p_k = k^{-2.2}$ (штрихована лінія)

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

У процесі досліджень маємо можливість простежити за розвитком та структуризацією комп'ютерної мережі «BW-Star & Fox Net» у часі. Характеристики цієї мережі в різні часові проміжки наведені в табл. 1. Якщо на початку становлення мережа займала проміжне місце між масштабною з пуасонівським розподілом ступенів вершин та безмасштабною мережами (рис. 1), то з часом відбуваються ріст 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 по 2011 рік, що вказує на те, що топологічні властивості мережі вже перебувають у стійкому стаціонарному стані. 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 імовірності приєднання цих вершин p_k залишаються практично незмінними, забезпечуючи тим самим степеневий розподіл ступенів вершин $p_k = k^{-\gamma}$ з незмінним показником степеня γ.

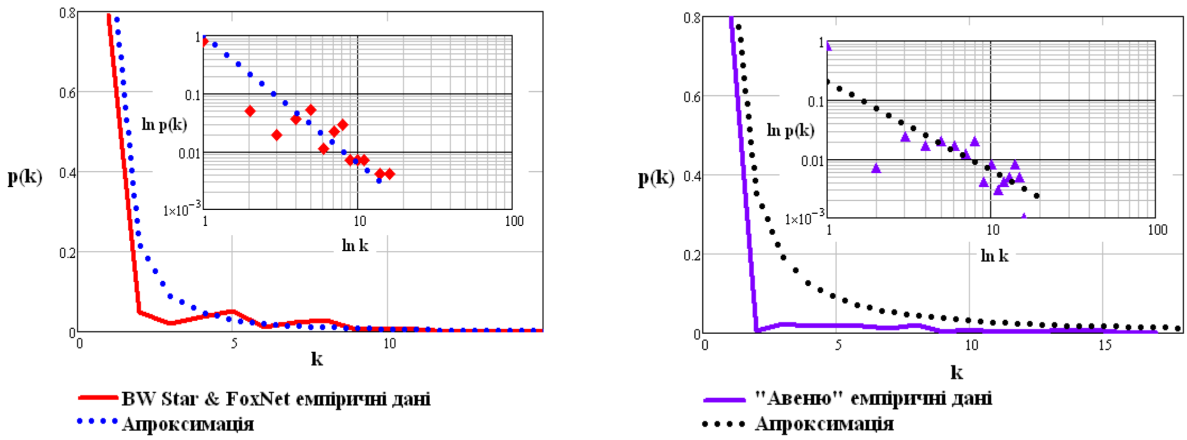

Рис. 2. Апроксимація «хвостів» розподілів ступенів вершин досліджуваних мереж

4. 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 «заготівок» (кінців) для майбутніх ребер;

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

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

• укладка графа здійснюється шляхом розміщення вершин з 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.

На основі розподілу p_k довільний граф може бути побудований $\prod_i k_i!$ різними способами, так як «заготівки» 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, так як з його допомогою можна побудувати мережу з довільним розподілом ступенів вершин.

Даний алгоритм реалізований у Processing, який є загальнодоступною мовою програмування і 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.

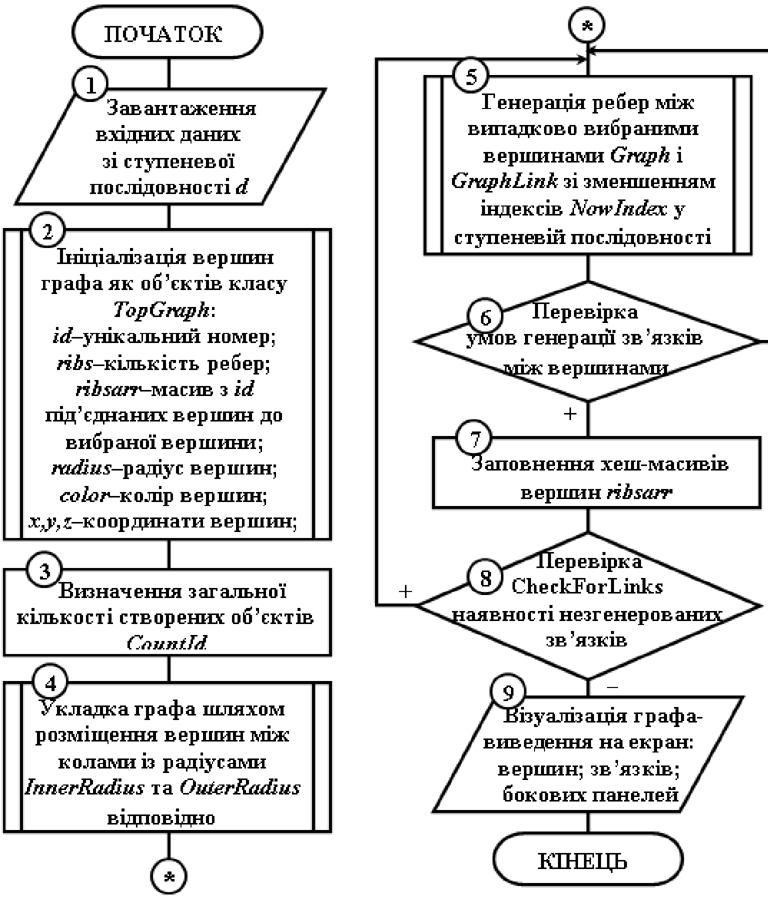

Рис. 3. Блок-схема програмної реалізації алгоритму моделювання

5. Комп'ютерний експеримент

Результатом програмної реалізації запропонованого алгоритму (рис. 3) є власне комп'ютерна мережа, зображена у вигляді стохастичного графа з відомим числом вершин і заданим розподілом імовірностей 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 в м. 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 здійснений процес моделювання мережі з подальшою можливістю порівняння результатів моделювання з характеристиками досліджуваних мереж, наведеними у табл. 1, та оцінювання адекватності опису моделлю реальної структури.

Провівши апроксимацію «хвостів» розподілів ступенів вершин, проілюстровану 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 імовірностей приєднання користувачів у мережу з відповідними γ. Як ілюстрація на рис. 4 наведені 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.

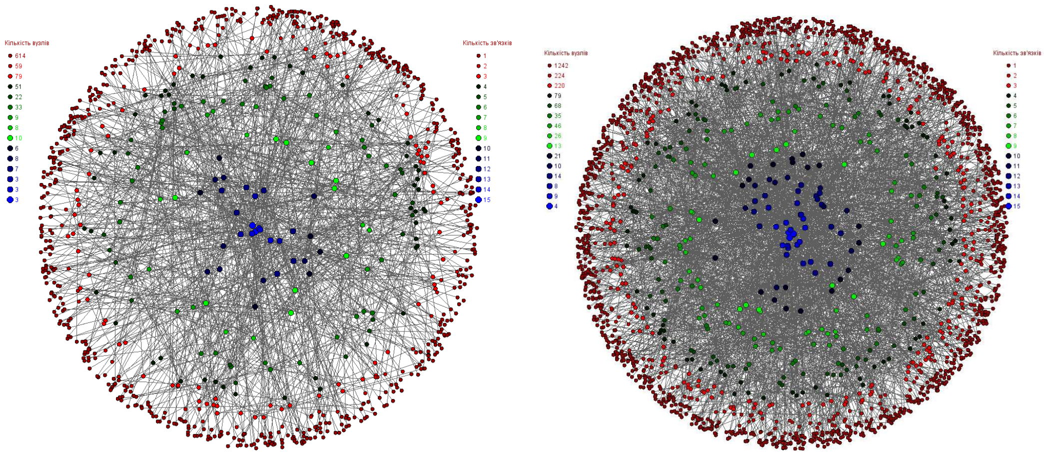

Рис. 4. Приклади графів, які відповідають мережам з різними законами залежності: $p_k = k^{-\gamma}$: (а) $\gamma = 2{,}4$, $N = 915$; (б) $\gamma = 2{,}1$, $N = 2023$

На рис. 4 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. 4 бачимо, що для малих значень параметра розподілу γ мережа кластеризується в один більший зв'язаний кластер, ніж 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 відмітити, що для переважної більшості комп'ютерних інтернет-мереж середній ступінь $\langle k \rangle$ в силу тих же причин прийматиме малі значення. Щодо параметра γ показника степеня степеневого розподілу, то його значення може бути різним. Більшим – у менш розгалужених систем 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. 4б), таких як мережа DSS-Group у м. 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.

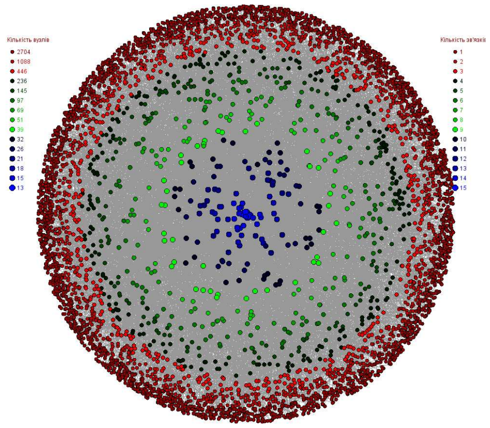

Рис. 5. Стохастичний граф для $p_k = k^{-2.2}$ при $N = 5000$

На рис. 5 зображений стохастичний граф, який відображає стан реальної комп’ютерної мережі з $p_k = k^{-2.2}$ в момент з кількістю зв’язків у ній $N = 5000$. Мережа достатньо структуризована, на периферії велика кількість вузлів з $k = 1$, які позначають користувачів. Запропонований алгоритм та його програмна реалізація дозволяють генерацію комплексних мереж з природною кількістю зв’язків у них $N \sim 10^3 \div 10^5$.

6. Прогнозовані результати розвитку мережі

Необхідно знати і вміти моделювати не тільки структуру зв’язків у даний момент часу, але і 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 завдяки щорічному приросту $\Delta N \approx 110$ 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 оцінки 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

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

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 [9]. Алгоритми моделювання, запропоновані в даній роботі, мають стати засобом для вироблення підходів до діагностики процесів розповсюдження комп'ютерних вірусів у комп'ютерних мережах та дослідження вразливості й стійкості останніх до спрямованих атак.

Аналіз реальних безмасштабних мереж WWW та Інтернету [10, 11], метаболізму [12], мережі харчування [13] 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.

7. Висновки

Завдяки дослідженням, вивчено підхід до моделювання динаміки розвитку та становлення комп'ютерних мереж з використанням апарату складних мереж. У рамках запропонованої схеми розроблене програмне забезпечення у середовищі Processing для моделювання комплексних комп'ютерних мереж.

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

Проведені оцінки, 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 рішення задачі про стійкість безмасштабних комп'ютерних мереж до спрямованих атак та розповсюдження вірусів у них.

СПИСОК ЛІТЕРАТУРИ

1. Нікольський Ю.В. Дискретна математика / Нікольський Ю.В., Пасічник В.В., Щербина Ю.М. – Львів: «Магнолія – 2006», 2009. – 432 с.
2. Newman M.E.J. The Structure and Function of Complex Networks / M.E.J. Newman // SIAM Review. – 2003. – Vol. 45, N 2. – P. 167 – 256.
3. Erdős P. On the evolution of random graphs / 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 / O. Frank, D. Strauss // Journal of the American Statistical Association. – 1986. – Vol. 81. – P. 832 – 842.
5. Watts D.J. Collective dynamics of "small-world" networks / D.J. Watts, S.H. Strogatz // Nature. – 1998. – Vol. 393. – P. 440 – 442.
6. Barabasi A.-L. Emergence of scaling in random networks / A.-L. Barabasi, R. Albert // Science. – 1999. – Vol. 286. – P. 509 – 512.
7. Price D.J. de S. A general theory of bibliometric and other cumulative advantage processes / D.J. de S. Price // Journal of the American Society for Information Science. – 1976. – Vol. 27. – P. 292 – 306.
8. Складні мережі / Ю. Головач, О. Олемський, К. фон Фербер [та ін.] // Журнал фізичних досліджень. – 2006. – Т. 10, № 4. – С. 247 – 289.

9. Stochastic simulation of HIV population dynamics through complex network modeling / P.M.A. Sloot, S.V. Ivanov, A.V. Boukhanovsky [et al.] // International Journal of Computer Mathematics. – 2008. – Vol. 85, N 8. – P. 1175 – 1187.
10. Albert R. Error and attack tolerance of complex networks / R. Albert, H. Jeong, A.-L. Barabasi // Nature (London). – 2000. – Vol. 406. – P. 378 – 381.
11. Tu Y. How robust is the Internet? / Y. Tu // Nature (London). – 2000. – Vol. 406. – P. 353 – 354.
12. The large-scale organization of metabolic networks / H. Jeong, B. Tombor, R. Albert [et al.] // Nature (London). – 2000. – Vol. 407. – P. 651 – 654.
13. Sole R.V. Complexity and fragility in ecological networks / R.V. Sole, J.M. Montoya // Proc. R. Soc. Lond. – 2001. – B 268. – P. 2039 – 2045.

Стаття надійшла до редакції 27.04.2012