

## КЛАСТЕРНИЙ АНАЛІЗ ЯК МЕТОДОЛОГІЧНИЙ ІНСТРУМЕНТАРІЙ ДОСЛІДЖЕННЯ БІБЛІОТЕК

© Кунанець Н.Е., Камінський Р.М., 2014

Наведено просту технологію ієрархічного агломеративного кластерного аналізу 20 бібліотек різного типу, поданих вибірками класифікаційних ознак однакових обсягів. Ця технологія реалізована в середовищі MsExcel-2003. Вона передбачає: створення таблиці “об’єкт-властивість”, побудову матриці близькостей, визначення структури дендрограми та інтерпретації кластерів.

Ключові слова: кластерний аналіз, таблиця “об’єкт–властивість”, матриця близькостей, дендрограма, кластери.

This article describes the simple technology of the hierarchical agglomerative cluster analysis of 20 different libraries, presented by the samples of classification attributes of the same volume. This technology is implemented in an environment of MsExcel-2003. It includes the transformation of one-dimensional data into multi-dimensional indexes, using descriptive statistics and distributions of individual parameters, the creation of the "object-property" table, the building of the proximity matrix, the definition of the dendrogram structure and the interpretation of the clusters.

Key words: the cluster analysis, the object-property table, the proximity matrix, dendrogram, clusters.

### Вступ

Вирішальну роль у забезпеченні досягнення значущого результату наукового дослідження відіграє оптимізація добору методів, які становлять базовий арсенал вибраного інструментарію, визначення можливої сфери використання кожного методу і відбору найефективнішого з них у кожному конкретному випадку. Методологія наукової діяльності – це процес, який передбачає застосування цілісної сукупності конкретних прийомів, підходів, способів, дій, спрямованих на отримання нових наукових результатів, досягнення поставленої мети та виконання запланованих завдань. Логічна побудова дослідницької роботи повинна ґрунтуватися на міждисциплінарному підході до вивчення бібліотек у контексті соціокомунікаційних процесів.

Бібліотека є інтегративним соціальним інститутом, тому багатогранною має бути і наука про бібліотеку, що зумовило вибір у цій роботі методів дослідження не лише з арсеналу бібліотекознавства, а й інформаційних технологій. І це цілком природно, оскільки сьогодні методологія наукових досліджень, поряд з використанням перевірених часом традиційних методів і підходів, потребує нових парадигм та концептуальних засад.

Бібліотека як система не є ізольованою від зовнішнього середовища, а пов’язана з ним цілою мережею зв’язків, через яку здійснює певний вплив, реалізуючи своє призначення, досягаючи у своєму функціонуванні певної мети. Крім цього, існують зв’язки іншого типу, які задають впливи та дії з боку зовнішнього середовища на бібліотеку. Ці зв’язки відіграють важливу роль у дослідженні процесів обміну інформацією між системою та зовнішнім середовищем, а також між елементами системи, які забезпечують реалізацію основних функцій системи. Дослідження та аналіз зв’язків дає змогу пізнавати об’єкти не безпосередньо, а опосередковано, через інші об’єкти, що перебувають з ними в тому чи іншому співвідношенні чи зв’язку.

Практично у кожній бібліотеці відтворюються найсуттєвіші зв’язки між структурними елементами та підсистемами, які неістотно змінюються у разі змін у системі та забезпечують її існування системи загалом, і підтримку найважливіших її властивостей.

Без освоєння історичного досвіду були б неможливі поглиблені сучасні теоретичні розроблення. Опрацювання значного масиву праць провідних науковців – як бібліотекознавців, так і дослідників із суміжних галузей наук дало змогу сформулювати характеристичні ознаки низки бібліотек з метою проведення їх типізації та точніше здійснити їхній поділ на однорідні в цьому сенсі класи, використовуючи метод ієрархічного агломеративного кластерного поділу.

Правомочність такого дослідження підтверджується тим, що ще в 70-ті роки ХХ ст. американський вчений-дослідник Дж. Солтон запропонував використати метод кластерного аналізу для розподілу бібліотечного фонду [1]. На жаль, в українському бібліотекознавстві інструментарій кластерного аналізу для дослідження бібліотек не застосовують.

**Мета статті:** проаналізувати можливості застосування методу ієрархічного кластерного аналізу у бібліотекознавстві, зокрема в задачах усебічного дослідження бібліотек, та показати його реалізацію за допомогою офісного пакета Ms Excel для поділу множини бібліотек на однорідні групи за принципом подібності на підставі їх класифікаційних характеристик, поданих таблицею “об’єкт-властивість”.

### 1. Суть кластерного аналізу

Перше застосування кластерний аналіз як метод знайшов у галузі соціології. Назва “кластерний аналіз” походить від англійського слова cluster – гроно, скупчення. Вперше в 1939 р. визначив предмет кластерного аналізу і подав його опис дослідник Р. Тріон [2]. Кластерний аналіз – це статистична процедура, що виконує багатовимірну класифікацію даних, які містять інформацію про вибірку об’єктів, тобто впорядковує об’єкти в порівняно однорідні групи. Тим самим розв’язується задача класифікації даних з використанням чітко сформованого математичного апарату.

Суть цієї процедури полягає в тому, що об’єкти, які необхідно класифікувати, подаються тим самим вектором (набором) індивідуальних ознак цих об’єктів у формі таблиці “об’єкт-властивість”, на підставі якої будується матриця відстаней (подібностей, близькостей), за якою і здійснюється кластеризація. Зміст кластеризації полягає в тому, що об’єкти, які за цими ознаками є найближчими один до одного, утворюють окрему групу – кластер. Основним критерієм зарахування об’єктів до тієї чи іншої групи, тобто проведення кластеризації, є у певний спосіб подана та виміряна “відстань” між об’єктами.

Методи кластерного аналізу можна застосовувати в різних випадках, зокрема йдеться про просте групування, у якому усе зводиться до утворення груп за кількісною ознаковою подібністю. Залежно від конкретної прикладної задачі мета кластерного аналізу може бути різною, наприклад:

- зрозуміти структуру множини об’єктів, розбивши їх на однорідні, в тому чи іншому сенсі, групи і, тим самим, спростити подальшу обробку даних для прийняття рішень, працюючи з кожним кластером окремо;

- виділити нетипові об’єкти, які не належать до жодного з кластерів. Цю задачу називають однокласовою класифікацією виявлення нетиповості або новизни;

- зменшити, у випадку надвеликих вибірок  $X$ , обсяг даних для збереження, залишивши по одному, найхарактернішому представнику з кожного кластера;

- дослідити динаміку об’єктів у процесі їхньої експлуатації за зміною відстаней всередині класів та між класами.

Методи кластерного аналізу можна застосовувати навіть тоді, коли необхідно здійснити звичайний поділ множини об’єктів на групи лише за кількісною подібністю вибраних ознак.

Алгоритми кластерного аналізу мають розроблену програмну реалізацію, що дозволяє ефективно розв’язувати задачі великої розмірності, проте такі програмні засоби є ліцензованими, а отже, платними та малодоступними. Тому його реалізація в середовищі Ms Excel є вельми актуальною задачею.

Метод ієрархічного агломеративного кластерного аналізу (який інколи називають чисельною таксономією) наявних даних здійснює класифікацію, яка раніше не існувала, або, переглядаючи дані, знову створює нову, ігноруючи попередню.

Результат проведення кластерного аналізу подається спеціальним графіком – дендрограмою, яка в натуральному масштабі відображає близькість або подібність між об'єктами, що класифікуються.

На підставі дендрограми можна виділити різні за обсягом кластери, причому відстані між об'єктами в кластері завжди будуть меншими, аніж найменша відстань між об'єктами з двох різних кластерів.

У цьому дослідженні з використанням методу ієрархічного агломеративного кластерного аналізу проведено класифікацію бібліотек за вираними ознаками й критеріями. Головне призначення застосування кластерного аналізу в нашому дослідженні – це підтвердження коректності та правильності класичного підходу до розподілу бібліотек за типами. Основною властивістю цих типів є те, що бібліотеки, які в результаті кластерного аналізу належать одному кластеру, за визначеними ознаками подібніші між собою, аніж з об'єктами, зарахованими до решти кластерів.

Для кластерного аналізу дані, зібрані в процесі дослідження чи отримані експериментально, подають у формі таблиці “об’єкт–властивість”, в той спосіб, що першим стовпчиком є назви об’єктів, які підлягають групуванню, а решта стовпчиків відповідають конкретним властивостям (ознакам, характеристикам, показникам тощо) і містять їх конкретні значення.

Його задача полягає в тому, щоб на підставі даних, поданих таблицею “об’єкт–властивість”, розділити множину  $G = \{g_i : g \in G, i = 1, 2, \mathbf{K}, m\}$  цих об’єктів на  $k$  ( $k$  – ціле число) кластерів – неперетинних підмножин  $Q_1, Q_2, \mathbf{K}, Q_k$  так, щоб їх об’єднання відповідало усій

множині  $\bigcup_{j=1}^k Q_j = G$ , а їх перетин був порожньою множиною  $\bigcap_{j=1}^k Q_j = \emptyset$ . Інакше кажучи:

- а) кожний об’єкт  $g_i$  має належати одній і тільки одній підмножині розбиття;
- б) об’єкти, що належать тому самому кластеру, були подібними;
- в) об’єкти, що належать до різних кластерів, були відмінними.

За отриманими даними має бути побудована дендрограма і наведена її інтерпретація стосовно варіантів вибору кількості кластерів.

Для цього в дослідженні використано методологію ієрархічного кластерного аналізу [3]. Реалізація методології цього виду кластерного аналізу здійснена безпосередньо в середовищі табличного процесора Ms Excel без залучення будь-яких додаткових підпрограм і для типологічного поділу бібліотек передбачає такі кроки.

## 2. Формування таблиці “об’єкт–властивість”

Для поділу множини відібраних об’єктів на кластери, в кожний з яких входять об’єкти, найвідповідніші йому за типом, необхідно насамперед створити таблицю “об’єкт–властивість”, в якій об’єктами є бібліотеки, а властивостям відповідають індивідуальні та однакової розмірності вектори (набори) характеристичних класифікаційних ознак. Для цього необхідно:

- 1) утворити множину бібліотек, які є об’єктами кластеризації;
- 2) визначити для кожної бібліотеки множину класифікаційних ознак та її потужність (обсяг) в тій самій номінальній шкалі;
- 3) вибрати шкалу значень класифікаційних ознак;
- 4) сформувати таблицю “об’єкт–властивість”.

Виокремлена нами на основі класичних критеріїв типологічного поділу множина об’єктів-бібліотек з яскраво вираженими типологічними ознаками подана в табл. 1.

Розподіл бібліотек здійснюється за низкою класифікаційних ознак. При цьому кластерний аналіз не накладає жодних обмежень на вид об’єктів, але кожен об’єкт має бути поданий тим самим набором ознак. Саме формування такого набору ознак-характеристик бібліотек є одним із найважливіших завдань дослідження.

Реалізація цього кроку полягає у визначенні набору ознак, які найкраще та найповніше відображають особливості бібліотек. В результаті проведеного відбору на основі аналізу та

опрацювання доступного первинного інформаційного матеріалу таким набором є множина ознак  $X = \{x_1, x_2, \mathbf{K}, x_m\}$ , де  $m = 8$ . В табл. 2 наведені: номінали класифікаційних ознак, їх характеристики, індивідуальні значення шкал кількісного оцінювання класифікаційних ознак.

Таблиця 1

### Множина вибраних бібліотек

№ з/п	Назва бібліотеки	Примітка
1	Почаївської лаври	
2	Ставропігійського братства	
3	Наукового товариства ім. Т. Шевченка	
4	Просвіти	Із читальнями
5	Наукового богословського товариства	
6	Союзу українок	
7	Рідної школи	Із шкільними
8	Національного музею	
9	Музею Дідушицьких	
10	Студіону	
11	Народного дому	
12	Ставропігійського інституту	
13	Закладу Осолінських	
14	Міська публічна	
15	Військова	
16	Архіву	
17	Львівського університету	Враховуючи кафедральні
18	Львівської політехніки	Враховуючи кафедральні
19	Аграрної академії	Враховуючи кафедральні
20	Богословської академії	Враховуючи кафедральні

Таблиця 2

### Шкали атрибутів бібліотек

№ з/п	Назва атрибуту	Характеристика	Шкала	Примітка
1	Мотивація заснування – $x_1$	Кількість	10	Кількість мотивів (обґрунтувань)
2	Функціональні завдання – $x_2$	Кількість	10	
3	Походження фондів – $x_3$	Кількість	2	За власний кошт 0 або акція ззовні 1
4	Приміщення для фондів – $x_4$	Кількість	8	Територіально розділені кімнати
5	Критерії комплектування – $x_5$	Кількість	2	Існують 1 або відсутні 0
6	Тематичне наповнення – $x_6$	Кількість	12	Вид тематики
7	Зв'язок з особистостями – $x_7$	Кількість	10	
8	Категорії читачів – $x_8$	Кількість	10	

Вибрані для кластеризації ознаки об'єктів подані кількісними значеннями, отриманими в результаті експертного оцінювання за наведеними відповідними шкалами, значення яких подані в балах. Бали визначено на основі аналізу джерельної бази з історії досліджуваних бібліотек [4].

Два з поданих атрибутів, а саме “Походження фондів” та “Критерії комплектування”, представлені в бінарному вимірі, решта за визначеною та встановленою на основі дослідження шкалою.

Оцінювання значень ознак вибраних нами бібліотек подано у табл. 3 у формі таблиці “об'єкт–властивість”, а точніше – таблиці “бібліотека–класифікаційні ознаки”.

Таблиця 3

Таблиця “бібліотека–класифікаційні ознаки”

№ з/п	Назва бібліотеки	Номери назв класифікаційних ознак							
		1	2	3	4	5	6	7	8
1	Монастирів	6	5	0	3	0	4	2	2
2	Ставропігійського братства	3	6	0	4	0	3	2	3
3	Наукового товариства ім. Т. Шевченка	3	4	1	2	1	7	5	4
4	Просвіти	5	4	0	1	1	6	8	4
5	Наукового Богословського. товариства	4	4	0	3	1	5	5	3
6	Союзу українок	4	6	1	2	0	4	9	4
7	Рідної школи	7	9	1	5	1	10	6	6
8	Національного музею	6	5	1	4	1	6	7	7
9	Музею Дідушицьких	8	4	0	3	1	7	6	5
10	Студіону	3	3	0	2	0	4	4	2
11	Народного дому	9	6	1	3	1	6	5	3
12	Ставропігії	4	2	0	2	0	6	3	3
13	Закладу Осолінських	5	4	0	6	1	11	7	8
14	Міської публічної	10	7	1	8	1	12	5	8
15	Військової	4	6	1	4	1	10	3	3
16	Архіву	3	8	1	6	1	9	2	5
17	Львівського університету	10	10	1	8	1	12	8	5
18	Львівської політехніки	10	10	1	8	1	7	5	5
19	Аграрної академії	7	5	1	2	1	3	2	5
20	Богословської Академії	5	3	1	1	1	2	3	2

### 3. Побудова матриці близькостей

Подібність або відмінність між об'єктами, що класифікуються, встановлюється залежно від метричної відстані між ними. Якщо кожен об'єкт описується  $k$  ознаками, то він може бути поданий як точка в  $k$ -вимірному просторі, і схожість з іншими об'єктами визначатиметься як відповідна відстань, обчислена за цими  $k$  ознаками. Кількісна оцінка подібності здійснюється на основі поняття метрики.

Для виконання кластерного аналізу необхідно побудувати матрицю близькостей, використовуючи для цього вже таблицю “бібліотека–класифікаційні ознаки”. Матрицю близькостей будуть за допомогою відповідних метрик, які визначають відстань між об'єктами, використовуючи значення наборів їх ознак. Ця процедура потребує вибору відповідної метрики для обчислення відстаней – близькості між кожною бібліотекою і рештою бібліотек.

Однією з найвикористовуваніших метрик у кластерному аналізі є евклідова відстань [3], оскільки вона відповідає інтуїтивним уявленням про близькість і своєю квадратичною формою описує класичні статистичні конструкції.

Проте евклідова відстань передбачає використання кількісних значень ознак або попереднього нормування їх величин. Тому в цьому дослідженні для побудови матриці близькостей використано мангеттенську метрику, яка найбільше відповідає фізичній суті процедури визначення відстаней між об'єктами, поданими бальними шкалами.

Відстанню (метрикою) між об'єктами в просторі параметрів вважатимемо таку величину  $d_{ab}$ , яка задовольняє аксіоми:

$$A1. d_{ab} > 0, d_{ab} = 0; \quad A2. d_{ab} = d_{ba}; \quad A3. d_{ab} + d_{bc} \geq d_{ac}.$$

Для побудови матриці у нашому дослідженні вибрано саме мангеттенську метрику, яка відповідає евклідовій метриці порядку один. За фізичним змістом об'єктів, що класифікуються, мангеттенська метрика відповідає цим аксіомам. Отже, відстань  $d_{ij}$  між  $i$ -м і  $j$ -м об'єктами визначається як сума абсолютних значень різниць між двома однойменними ознаками цих об'єктів. Аналітично це виглядає так:

$$d_{ij} = |x_{1i} - x_{1j}| + |x_{2i} - x_{2j}| + \dots + |x_{8i} - x_{8j}| = \sum_{k=1}^8 |x_{ki} - x_{kj}|.$$

Використовуючи обчислювальні можливості офісного пакета Ms Excel, а саме табулювання функцій методом автозаповнення та фіксації назви стовпчика та номера рядка, та одночасне зчитування значень класифікаційних ознак з таблиці "бібліотека-класифікаційні ознаки" та її поруч на листі розташованій копії, за  $N = 20$  ітерацій отримуємо матрицю близькостей у вигляді табл. 4.

Таблиця 4

Матриця близькостей

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII	XIX	XX
I	0	7	15	15	9	14	23	15	14	8	12	10	24	30	14	19	33	25	8	10
II	7	0	14	18	10	13	24	18	19	9	15	11	25	31	11	14	34	26	11	13
III	15	14	0	8	6	11	18	12	9	9	11	9	17	25	11	14	28	20	13	13
IV	15	18	8	0	8	9	20	10	9	13	13	11	15	29	17	22	26	24	15	13
V	9	10	6	8	0	11	20	12	9	7	9	7	17	27	11	18	30	22	13	11
VI	14	13	11	9	11	0	21	13	16	12	14	14	22	30	16	21	27	25	14	16
VII	23	24	18	20	20	21	0	12	13	27	15	25	13	13	13	12	12	12	19	27
VIII	15	18	12	10	12	13	12	0	9	19	11	17	11	19	15	18	22	18	13	19
IX	14	19	9	9	9	16	13	9	0	16	8	14	14	20	16	19	21	15	12	18
X	8	9	9	13	7	12	27	19	16	0	16	6	24	34	16	21	37	29	14	8
XI	12	15	11	13	9	14	15	11	8	16	0	14	22	18	12	19	21	13	12	16
XII	10	11	9	11	7	14	25	17	14	6	14	0	22	32	12	19	35	27	14	10
XIII	24	25	17	15	17	22	13	11	14	24	22	22	0	14	16	17	19	23	24	26
XIV	30	31	25	29	27	30	13	19	20	34	18	32	14	0	20	19	9	11	26	34
XV	14	11	11	17	11	16	13	15	16	16	12	12	16	20	0	9	23	21	16	16
XVI	19	14	14	22	18	21	12	18	19	21	19	19	17	19	9	0	20	16	17	23
XVII	33	34	28	26	30	27	12	22	21	37	21	35	19	9	23	20	0	8	29	37
XVIII	25	26	20	24	22	25	12	18	15	29	13	27	23	11	21	16	8	0	21	29
XIX	8	11	13	15	13	14	19	13	12	14	12	14	24	26	16	17	29	21	0	10
XX	10	13	13	13	11	16	27	19	18	8	16	10	26	34	16	23	37	29	10	0

Ця матриця є симетричною відносно діагоналі, значення на якій є нулями. Рівність нулю означає відстань між тим самим об'єктом, яка, природно, дорівнює нулю.

#### 4. Побудова дендрограми

Завершальною процедурою ієрархічного агломеративного кластерного аналізу є визначення параметрів дендрограми, її побудова та інтерпретація щодо вибору кластерів та їх кількості. На підставі таблиці близькостей і вибраної стратегії об'єднання, суть якої полягає в тому, що об'єкти в таблиці поступово, крок за кроком, об'єднуються в групи-кластери. Ієрархічний агломеративний кластерний аналіз для роботи з матрицею близькостей використовує такі стратегії: найближчого сусіда, дальнього сусіда, групового середнього, центроїдну стратегію, основу на прирості суми квадратів, гнучку [3].

Початкові дії в процедурах ієрархічного агломеративного кластерного аналізу однакові. Спочатку серед всіх об'єктів матриці близькості шукають пару об'єктів з найменшою мірою відстані й об'єднують їх в одну групу. Стовпчики і рядки цих об'єктів вилучаються, а на їх місце вставляють новий стовпчик і рядок з перерахованими значеннями ознак, так, щоб не порушити діагональ нулів. У результаті розмір матриці близькостей зменшується на одиницю, а знайдене найменше значення стає параметром дендрограми, оскільки визначає відстань між цими об'єктами, а саму групу позначають номером  $n + 1$ . На кожному наступному кроці відбувається об'єднання або між двома об'єктами, або між об'єктом і групою, або між двома групами, для яких міра близькості мінімальна, здійснюють аналогічний перерахунок і об'єднані групи позначають  $n + 2$ ,  $n + 3$ , ...,  $n + (n - 1)$ . Завершується процедура тоді, коли розмірність матриці близькостей стане  $2 \times 2$ . Математичною підставою для об'єднання в групи є вираз Уїлліамса–Ланса [3]

$$d_{nk} = a_i \cdot d_{hi} + a_j \cdot d_{hj} + b \cdot d_{ij} + g \cdot |d_{hi} - dhj|, \quad (1)$$

де  $a_i, a_j, b, g$  – визначаються видом стратегії,  $d_{hi}$  і  $d_{hj}$  – стовпчики, які об'єднуються,  $d_{hk}$  – стовпчик в результаті об'єднання.

У цьому дослідженні використано гнучку стратегію, яка застосовується для будь-яких мір близькостей і визначається такими параметрами  $a_i = a_j = 0.625$ ,  $b = -0.25$  і  $g = 0$ . Процедура об'єднання за цією стратегією реалізується за таким алгоритмом:

$$d_{rs} = 0,625 \cdot (x_{rs} + x_{(r+1)s}) - 0,25(\min(x_{r(s-1)})), \quad (2)$$

де  $r$  – номер стовпчика,  $r = \overline{1, n}$ , а  $s$  – номер рядка,  $s = \overline{1, n}$ .

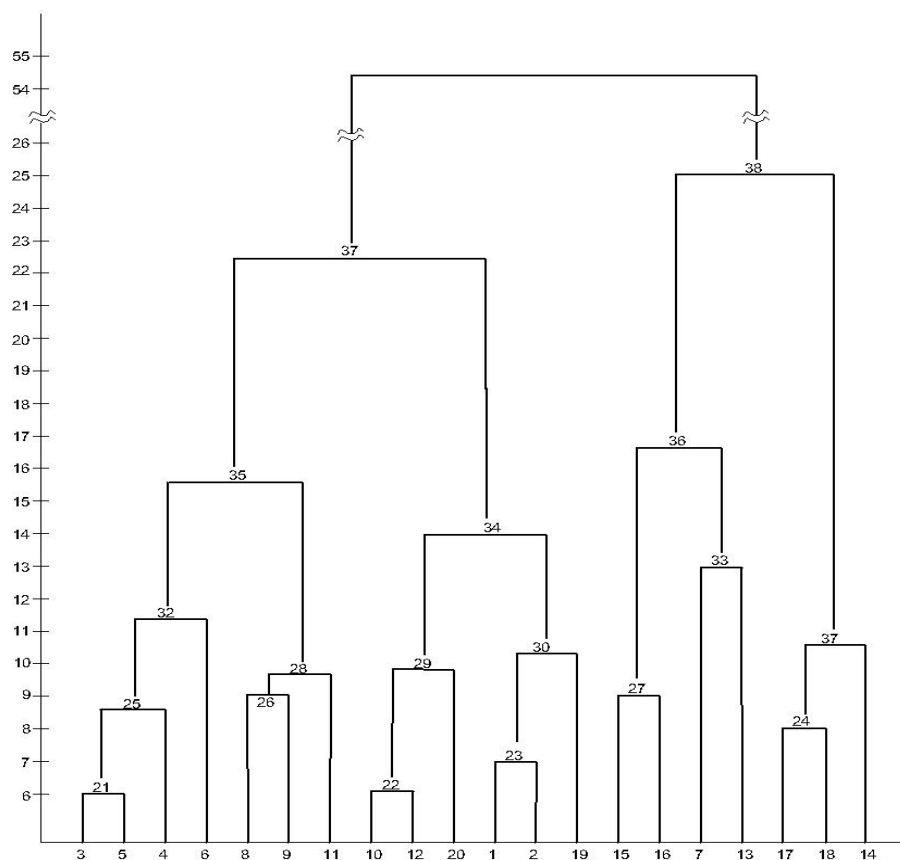
Процедура стратегії об'єднання полягає в тому, що знаходять у матриці близькостей найменше значення  $d_{rs}$ , яке дорівнює 6. Це дає можливість об'єднати об'єкти 3 і 5 в одну групу і присвоїти цій групі номер – 21. Для цієї групи виконують перерахунок стовпчика і рядка за виразом (2). Впродовж всієї роботи алгоритму для кожної нової групи здійснюється перерахунок за відповідним значенням груп, що об'єднуються, тобто вона має нові значення.

Далі знову шукаємо найменше значення – воно відповідає відстані між групами 10 і 12 і становить 6. Аналогічно здійснюємо перерахунок значень ознак і це об'єднання позначаємо номером 22, і так далі. Результати реалізації процедури об'єднання подано у табл. 5.

На жаль, графічне забезпечення Microsoft Excel не дає можливості безпосередньо побудувати дендрограму, оскільки її побудова спряжена не лише з масштабом відстаней, але і з місцем об'єктів, які тепер належать конкретним кластерам. Тому перший крок побудови ручний, а сам процес побудови починається з кореня дерева. Наприклад, номер 39 вказує на об'єднання двох груп 37 і 38, тобто можемо побудувати в довільному масштабі дві прямокутні гілки. Своєю чергою, номер 37 є групою з двох початкових об'єктів 34 і 35. Натомість вузол 38 об'єднує групи 36 і 31 групи, причому група 31 є об'єднанням груп 14-ї та 24-ї. Аналогічно аналізуємо решту груп. Процедuru повторюємо до побудови повного дерева, тобто поки не дійдемо до самих необ'єднаних об'єктів. Після побудови такого дерева стає зрозумілим ієрархічний порядок кластерів об'єктів, який означає: що вищий рівень (більша відстань між кластерами – вісь ординат), то більше об'єктів містять кластери. Далі проводять вертикальні відрізки відповідно до шкали відстаней по осі ординат. В результаті дендрограма має вигляд, зображений на рисунку.

## Покрокове об'єднання кластерів

ЗНАЧЕННЯ ДЕНДРОГРАМИ		
Об'єднання груп	Відстані між групами, що об'єднуються	Новоутворені групи
d 3-5 =	6	21
d 10-12 =	6	22
d 1-2 =	7	23
d 17-18 =	8	24
d 21-4 =	8,5	25
d 8-9 =	9	26
d 15-16 =	9	27
d 26-11 =	9,63	28
d 22-20 =	9,75	29
d 23-19 =	10,13	30
d 14-24 =	10,5	31
d 25-6 =	11,16	32
d 7-13 =	13	33
d 30-29 =	13,89	34
d 32-28 =	15,43	35
d 33-27 =	16,59	36
d 34-35 =	22,51	37
d 36-31 =	24,94	38
d 37-38 =	54,67	39



Дендрограма кластеризації бібліотек



## Висновок

Отримані результати кластерного аналізу не лише підтверджують класичну схему класифікації досліджуваних бібліотек, спрощують саму процедуру поділу, але й можуть використовуватись для подальших теоретичних розробок класифікаційних схем бібліотек.

Ми провели в “ручному режимі” ієрархічний агломеративний кластерний аналіз лише для невеликої вибірки бібліотек. Якщо їх кількість велика (тисячі і десятки тисяч елементів), вручну реалізувати таку технологію недоцільно і неможливо. Для цього мають бути використані відповідні програмні засоби, поглиблений аналіз результатів кластеризації великої кількості бібліотек може слугувати основою для побудови відповідної онтології у галузі бібліотекознавства.

Застосування методу кластеризації може надалі орієнтуватися на створення алгоритмів ефективного інформаційного пошуку, а також побудову науково обґрунтованих класифікаційних систем, зорієнтованих на галузь бібліотекознавства. Метод ієрархічного агломеративного кластерного аналізу може застосовуватися і для типологічного чи змістового розподілу фондів бібліотек, дослідження тематичного та видового їх складу. Цей метод можна вважати універсальним, таким, що дає можливість формалізувати типологічний поділ будь-яких об’єктів бібліотекознавства.

1. Солтон Дж. Динамические библиотечно-информационные системы / Дж. Солтис. – М.: Мир, 1979. – С. 9, 35, 369. 2. Tryon R.C. Cluster analysis / Tryon R.C. – London: Ann Arbor Edwards Bros, 1939. – 139 p. 3. Статистический метод для ЭВМ / пер. с англ. – М.: Наука, 1986. – С.269–301. 4. Кунанець Н. Е. Бібліотечна справа на західноукраїнських землях (XIII ст. – 1939 р.): консолідований інформаційний ресурс [Текст] : монографія / Н. Е. Кунанець. – Л.: Вид-во Львів. політехніки, 2011. – 448 с.

УДК 004.9

А.В. Катренко, А.С. Магац

Національний університет “Львівська політехніка”,  
кафедра інформаційних систем та мереж

## АНАЛІЗ МАТЕМАТИЧНИХ МОДЕЛЕЙ ПЛАНУВАННЯ В МУЛЬТИПРОЕКТНОМУ СЕРЕДОВИЩІ

© Катренко А.В., Магац А.С., 2014

Розглянуто проблематику організації, планування та підтримання управління в організаціях, які реалізують свою діяльність у вигляді проектів. Здійснено критичний аналіз функціонування у такому середовищі, вказано основні недоліки та методи їх подолання.

**Ключові слова:** мультипроект, оптимізація, офіс управління проектами, планування.

**Problems of organization, planning and maintenance management in organizations that implement their activity in the form of projects are considered. The critical analysis of the operation in this environment was carried out, major shortcomings and methods of overcoming them are pointed out.**

**Key words:** multiproject, optimization, project management office planning.

### Вступ. Обґрунтування актуальності проблеми

Сьогодні існує проблема у галузі управління проектами, оскільки незначна їх частина закінчується у терміни і обсяги, які передбачалися. Проблема полягає в тому, що багато організацій не знають, якими ресурсами вони реально володіють. Потрібно виокремлювати ключові ресурси. Часто буває, що кількість проектів є дуже великою і корисно відкинути частину з них.