

5. Doganis, R. Traffic forecasting and the gravity model [Text] / R. Doganis // Flight International. – 1966. – P.547–549.
6. Погребняк, Е. Б. Анализ методов формирования матрицы корреспонденций транспортной сети города [Текст] / Е. Б. Погребняк, Н. И. Самойленко // Коммунальное хозяйство городов. – Харьковская национальная академия городского хозяйства, 2006. – № 69. – С. 121-126.
7. Пархоменко, Л. О. Розроблення моделі прогнозування кореспонденцій пасажирів в умовах впровадження залізничного швидкісного пасажирського сполучення на основі нечітких реляційних обчислень [Текст] / Л. О. Пархоменко // Збірник наукових праць УкрДАЗТ. – 2012. – Вип. 131. – С. 109-115.
8. Рутковская, Д. Нейронные сети, генетические алгоритмы и нечеткие системы [Текст] / Д. Рутковская, М. Пилинский, Л. Рутковский; пер.с польск. И. Д. Рудинского. – М.: Горячая линия. – Телеком, 2004. – 452 с.
9. Peeva, K. Fuzzy Relational Calculus Theory: Applications and Software [Text] / Ketty Peeva, Yordan Kyosev // World Scientific Publishing Co. Pte. Ltd. – 2005. – 291 p.
10. Wright, A. "Genetic algorithms for real parameter optimization"[Text] / A. Wright // Foundations of Genetic Algorithms. – 1991. – V. 1. – P. 205-218.
11. Ротштейн, А. П. Идентификация нелинейной зависимости нечеткой базой знаний с нечеткой обучающей выборкой [Текст]/ А. П. Ротштейн, С. Д. Штовба // Кибернетика и системный анализ. – 2006. – No2. – С. 17-24.
12. Кисляков, А. В. Генетические алгоритмы: операторы скрещивания и мутации [Текст] / А. В. Кисляков // Информационные технологии. – 2001. – №1. – С. 29-34.

*Виконано побудову математичних моделей для задач кількісного оцінювання груп ризику інфікування вірусом імунодефіциту людини. Проведено аналіз технологій ідентифікації та кластеризації для розв'язання задачі оцінювання кількості представників груп ризику та задачі визначення міри належності особи до такої групи*

*Ключові слова: група ризику інфікування вірусом імунодефіциту людини, задача кластеризації, задача ідентифікації*

*Выполнено построение математических моделей для задач количественной оценки групп риска инфицирования вирусом иммунодефицита человека. Проведен анализ технологий идентификации и кластеризации для решения задачи оценки числа представителей групп риска и задачи определения степени принадлежности лица к такой группе*

*Ключевые слова: группа риска инфицирования вирусом иммунодефицита человека, задача кластеризации, задача идентификации*

УДК 004.94

# ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ КІЛЬКІСНОГО ОЦІНЮВАННЯ ГРУП РИЗИКУ ІНФІКУВАННЯ ВІРУСОМ ІМУНОДЕФІЦИТУ ЛЮДИНИ

О. Ю. Мулеса

Викладач

Кафедра кібернетики і прикладної математики  
Ужгородський національний університет  
вул. Університетська, 14, м. Ужгород, Україна, 88000  
E-mail: mulesa.oksana@gmail.com

## 1. Вступ

Світова спільнота об'єднує свої зусилля для подолання епідемії СНІДу [1]. Так, Об'єднаною програмою Організації Об'єднаних Націй з ВІЛ/СНІДу було розроблено 10 конкретних цільових показників на 2015 рік для направлення сумісних дій, серед яких [1]:

1. Знизити передачу ВІЛ статевим шляхом на 50% до 2015 року.

2. Знизити передачу ВІЛ серед споживачів ін'єкційних наркотиків (СІН) на 50% до 2015 року.

3. Викоринити нові випадки ВІЛ-інфікування серед дітей до 2015 року та значно знизити рівень материнської смертності внаслідок СНІДУ тощо.

Україна є однією з країн Східної Європи, найбільш охоплених епідемією ВІЛ-інфекції/СНІДу [2]. Згідно з останніми офіційними даними, масштаби епідемії ВІЛ-інфекції/СНІДу в Україні мають зростаючу динаміку, незважаючи на комплекс заходів з

протидії її поширенню, які реалізуються в країні. Для того, щоб можна було проводити стратегічне планування профілактичних заходів, подальше прогнозування рівня впливу соціальних програм, які впроваджуються на території України, необхідно визначити кількість представників груп високого ризику інфікування ВІЛ (ГВР).

## 2. Аналіз літературних даних та постановка проблеми

За останнє десятиліття в Україні здійснювалася не одна спроба оцінити чисельність груп ризику на національному та локальному рівнях. Основним результатом цих досліджень є узгоджені на національному та регіональних рівнях оцінки кількості представників ГВР, які рекомендовані для широкого використання [2].

Однак, зважаючи на специфіку цільових груп, необхідно пам'ятати про те, що всі оцінки мають розглядатися як не точні, оскільки всі вони базуються на певних припущеннях. Таким чином, усі оцінки чисельності груп ризику є гіпотетичними і не можуть розглядатися як однозначні та остаточні.

Загалом, суттєвою проблемою на сьогодні є те, що на даний момент не існує єдиних ustalених критеріїв віднесення до груп ризику. Основні критерії включення до груп ризику (період останньої практики, регулярність тощо) визначаються для кожного дослідження окремо. Так само не існує прийнятих критеріїв щодо розмежування груп ризику між собою, хоча відомо, що часто трапляються випадки «перетину сфер» ризикованої поведінки.

Традиційно, з метою проведення оцінки кількості представників ГВР, в різних містах України проводяться соціологічні дослідження, робляться спроби побудувати соціально-демографічний портрет особи – представника ГВР [3 – 5]. У роботах [2, 6] запропоновано підхід до оцінки кількості представників деяких ГВР, який базується на використанні статистичної інформації.

В основі даного підходу лежить використання методики RDS, теоретичною основою якої є теорія ланцюгів Маркова та теорія малих мереж, та методики TLS – вибіркового методу, відповідно до якого вибір учасників дослідження відбувається на спеціальних територіях (в місцях збору цільової групи) [7]. Застосування такого підходу із залученням міжнародних фондів спричиняє великі фінансові витрати, що не завжди є можливим та доцільним. До того ж, при такому підході можливі грубі узагальнення, які нівелюють вплив таких важливих чинників як релігійність населення того чи іншого регіону, рівень освіченості населення тощо.

Таким чином, провівши аналіз аспектів поширення ВІЛ-інфекції, можна зробити висновок про необхідність розробки методики оцінки кількості представників ГВР як на всій території України, так і на її окремих адміністративних частинах, в якій були б інтегровані як вже відомі підходи до розв'язання описаної задачі, так і нові підходи, які б дозволяли якнайповніше враховувати особливості вхідних даних, результати експертних

опитувань та професійну думку особи, що приймає рішення (ОПР).

## 3. Мета та задачі дослідження

Розглядаючи проблему оцінювання кількості представників ГВР, можна виділити дві взаємопов'язані задачі, які впливають одна з іншою [8]: задача визначення оціночної кількості представників ГВР серед дорослого населення окремо взятої території та задача визначення міри належності особи до ГВР.

Перша задача полягає в оцінюванні кількості представників ГВР у заданому регіоні. Для оцінювання є відомими соціально-демографічні показники населеного пункту та додаткові фактори, які, на думку ОПР, також впливають на результуючу оцінку: рівень релігійності населення, близькість до кордону, близькість до міжнародних автошляхів тощо. В такому представленні описану задачу відносять до задач кластеризації [9 – 11], які зводяться до розбиття множини об'єктів на групи найбільш подібних за деяким критерієм об'єктів.

У другій задачі необхідно визначити міру належності особи до ГВР. Для оцінювання надаються анкетні дані особи (вік, стать, освіта тощо) та дані, що характеризують місце проживання та місце діяльності особи.

В такій постановці дана задача може бути віднесена до задач ідентифікації [12], в основі яких лежить необхідність встановлення залежності

$$Y = F(X_1, X_2, \dots, X_m),$$

де  $(X_1, X_2, \dots, X_m)$  – набір вхідних параметрів,  $Y$  – деяка вихідна величина.

Також, в математичних моделях обох вказаних задач вхідними даними можуть бути результати соціологічних досліджень, соціально-демографічні показники тощо.

Попередній аналіз дозволяє зробити висновок про те, що через соціально-демографічний портрет особи можна визначити оціночну кількість осіб, що належать до ГВР, а потім для тих осіб, які входять в оціночну кількість визначити міру належності даній групі. І навпаки, визначивши для кожної особи міру можливої належності ГВР, можна оцінити кількість представників даної групи.

Тому задача оцінювання кількості осіб, що належать до ГВР може розглядатися одночасно як задача кластеризації і як задача ідентифікації параметрів системи.

Формальні постановки наведених задач можуть бути сконструйовані наступним чином.

**Задача кластеризації  $Z_1$ .** Нехай задано множину осіб  $O = \{O_1, O_2, \dots, O_n\}$ , які оцінюються за множиною критеріїв  $K = \{K_1, K_2, \dots, K_m\}$ . Необхідно в початковій множині виділити кластер осіб, які належать до ГВР.

**Задача ідентифікації  $Z_2$ .** Нехай задано особу  $O$ , яка характеризується вектором значень ознак  $X = (x_1, x_2, \dots, x_m)$ , та множину  $O' = \{O_1, O_2, \dots, O_n\}$  осіб, що належать кластеру, кожен елемент множини  $O'$ , в свою чергу, характеризується вектором значень ознак  $X^j = (x_{1j}, x_{2j}, \dots, x_{mj})$ ,  $j = \overline{1, N}$ .

На основі заданої інформації необхідно ідентифікувати залежність міри належності особи до кластера у від значень ознак

$$y = F(X_1, X_2, \dots, X_M)$$

і для даної особи визначити міру її належності кластеру.

#### 4. Методи розв'язання задачі кластеризації $Z_1$

При розв'язанні задач кластеризації ключовими є проблеми вибору метрики для визначення відстані між об'єктами та проблема врахування нечислових характеристик об'єктів.

Розглянемо деякі технології розв'язання задачі кластеризації  $Z_1$ .

Нейромережні технології. Задачу  $Z_1$  оцінки кількості представників певного кластера серед множини заданих об'єктів можна віднести до класу задач класифікації без учителя. Серед технологій, які широко застосовуються для розв'язання такого класу задач, можна виділити нейромережні технології навчання без учителя, розроблені Кохоненом [13] та іншими вченими.

Моделі такого типу не потребують цільового вектору, тобто не вимагають порівняння з наперед визначеними відповідями. Навчальний алгоритм в даних моделях настроює вагові коефіцієнти мережі таким чином, щоб отримати узгоджені вихідні вектори, тобто, щоб для достатньо близьких вхідних векторів мережа давала однакові виходи. Тому, у процесі навчання мережі виділяються статистичні властивості елементів навчальної множини і вхідні вектори групуються у кластери.

У загальному, НМ із шаром Кохонена для описаної задачі визначення кількості представників ГВР може бути представлена так як на рис. 1.

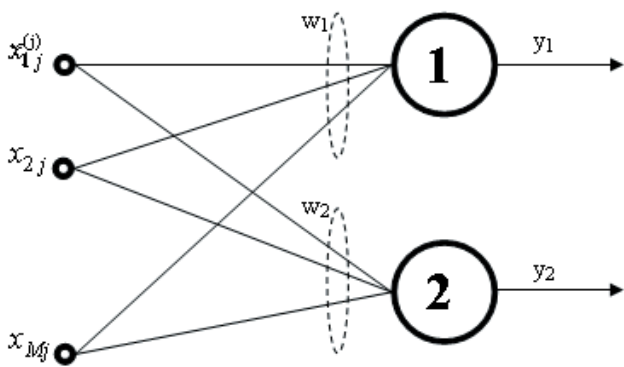


Рис. 1. Мережа Кохонена, що адаптована для розв'язання задачі кластеризації  $Z_1$

У даному випадку мережа буде складатись з двох нейронних елементів, тобто в результаті роботи алгоритму навчання НМ буде здійснено поділ вхідних образів на два кластери: кластер векторів, які належать особам, що відносяться до ГВР, та кластер векторів, які відповідають особам, що не належать до ГВР. Очевидно, що результат роботи алгоритму на-

вчання НМ залежить від початкових значень вагових коефіцієнтів нейронних елементів.

Представимо алгоритм функціонування НМ Кохонена для задачі оцінювання кількості осіб, що належать до ГВР, таким чином [13]:

Крок 1. Нормуємо вхідні вектори за одним із співвідношень

$$x'_{ij} = \frac{x_{ij} - \min_i x_{ij}}{\max_i x_{ij} - \min_i x_{ij}}, \quad x''_{ij} = \frac{\max_i x_{ij} - x_{ij}}{\max_i x_{ij} - \min_i x_{ij}}$$

Крок 2. Задаємо вагові коефіцієнти нейронних елементів шару Кохонена та нормуємо їх.

Крок 3. Для вхідного вектора визначаємо нейрон  $w^*$  за одним із співвідношень

$$d(X, w^*) = \min_{1 \leq l \leq k} d(X, w_l), \quad w^* \in \text{Arg max}_{1 \leq l \leq k} \sum_{i=1}^N x_i w_{li}$$

де  $k$  – кількість нейронів у мережі;  $X = (x_1, x_2, \dots, x_N)$  – вхідний вектор;  $w_l = (w_{l1}, w_{l2}, \dots, w_{lN})$  – вектор ваг  $l$ -го нейрону;  $d(a, b)$  – міра відстані між векторами.

Крок 4. Налаштовуємо ваги нейрона  $w^*$  за таким правилом:

$$w_1^{(k+1)} = w_1^{(k)} + \eta_1^{(k)} (X - w_1^{(k)}),$$

де  $k$  – номер циклу навчання НМ;  $\eta_1^{(k)}$  – коефіцієнт швидкості навчання  $l$ -го нейрону на  $k$ -му циклі навчання.

При цьому, на початкових етапах коефіцієнт швидкості навчання  $\eta_1^{(k)}$  фіксують на рівні 0,7 і поступово зменшують в процесі навчання НМ. На виході нейрона  $w^*$  одержуємо 1, а на виходах всіх інших нейронів – 0.

Крок 5. Якщо похибка мережі  $E = \frac{1}{N} \sum_{j=1}^N \|X^{(j)} - w^*\|$  – достатньо мала величина, то алгоритм закінчив роботу, інакше переходимо до кроку 3.

Метод навчання НМ Кохонена в такій інтерпретації дозволяє розділити вхідні вектори, що характеризують осіб із заданої множини, на два кластери близьких між собою векторів. Для того, щоб за допомогою даного алгоритму можна було отримати правильні розв'язки, необхідно, щоб виконувались дві наступні умови:

- було справедливим припущення про те, що всі особи, які належать до ГВР, характеризуються "схожими" соціально-демографічними портретами, тобто їх вектори значень ознак "близькі" між собою;
- початкове значення вектору вагових коефіцієнтів одного з нейронних елементів було "близьким" до векторів, що описують соціально-демографічний портрет особи, яка належить до ГВР.

Очевидно, що перше припущення не відповідає дійсності та його врахування призведе до отримання невірних результатів.

Що стосується початкових значень вагових коефіцієнтів, то їх вдале задання залежить лише від кваліфікації спеціаліста та його компетентності як в галузі обробки даних, так і в теорії нейронних мереж.

Для того, щоб отримати можливість ігнорувати першу умову та допустити можливість віднесення до заданого кластеру векторів, значення яких не є "близькими" між собою, необхідно умовно розділити кластер на складові частини "близьких" між собою елементів.

Для реалізації такого підходу при розв'язуванні задачі оцінювання кількості осіб, що належать до ГВР, необхідно виділити основні типи соціально-демографічного портрету осіб, що належать до вказаної групи.

Але, зважаючи на специфіку цільової групи дослідження та можливість входження до неї осіб з різними соціально-демографічними портретами, – задача виділення основних типів таких портретів може звестися до простого переліку всіх можливих соціально-демографічних портретів. Таким чином, до заданого кластеру увійдуть всі особи, що належать початковій множині.

Нечіткий метод кластеризації. Задачу  $Z_1$  визначення кількості осіб, що належать до ГВР, можна звести до розбиття початкової множини осіб на два кластери: кластер осіб, що належать заданій групі, та кластер осіб, що до неї не належать.

Проте, зважаючи на специфіку цільової групи дослідження, особливості (складність отримання, неточності в даних тощо), що можуть виникнути при отриманні інформації про конкретних осіб, доцільним є застосування методів кластеризації, які передбачають можливість віднесення одного і того ж об'єкта до декількох кластерів з різною мірою належності об'єкта цим кластерам.

Таким методом, який базується на припущенні, що один і той же об'єкт може з різною мірою належності належати декільком кластерам, є метод нечіткої кластеризації С-середніх [14, 15]. Для застосування цього методу при розв'язанні описаної задачі зробимо наступні підготовчі кроки.

Розглянемо задачу  $Z_1$ . На початковому етапі виділимо два кластери  $H_1, H_2$  – відповідно кластер осіб, що належать до ГВР, та кластер осіб, що до неї не належать.

Задамо параметр зупинки алгоритму  $\epsilon \in (0;1)$  та значення міри нечіткості  $m, 1 \leq m < \infty$  [16]. При  $m=1$  отримуємо чітке розбиття, тобто розбиття, у якому кожен об'єкт відноситься тільки до одного кластеру. Із збільшенням параметру  $m$  збільшується міра розмитості виділених кластерів.

Тоді в загальному випадку метод С-середніх для задачі визначення кількості осіб, що належать до ГВР, представимо як послідовність наступних кроків:

Крок 1. Для кожного об'єкта із заданої множини  $\bar{O} = \{O_1, O_2, \dots, O_n\}$  довільним чином задаємо міру належності кожному з кластерів та формуємо матрицю  $U^0 = (u_{ij})$ ,  $u_{ij} \in [0;1]$ , де  $u_{ij}$  – початкова міра належності  $i$ -го об'єкта  $j$ -му кластеру ( $i=1, n, j=1, 2$ ), де  $u_{i1} + u_{i2} = 1, \forall i \in \{1, \dots, n\}$ .

Крок 2. Обчислюємо центри кластерів за формулою

$$\bar{c}_j = \frac{\sum_{i=1}^n (u_{ij})^m \times X^{(i)}}{\sum_{i=1}^n (u_{ij})^m}.$$

Крок 3. Знаходимо відстані від об'єктів до центрів кластерів за формулою

$$d_{ij} = \sqrt{\|c_j - X^{(i)}\|^2}, \quad (i = \overline{1, n}, j = \overline{1, 2}).$$

Крок 4. Обчислюємо коефіцієнти матриці  $U$  таким чином:

якщо  $d_{ij} = 0$ , то  $u_{ij} = 1, u_{i,3-j} = 0$ ;

$$\text{якщо } d_{ij} > 0, \text{ то } u_{ij} = \frac{1}{\left(d_{ij}^2 \left(\frac{1}{d_{i1}^2} + \frac{1}{d_{i2}^2}\right)\right)^{\frac{1}{m-1}}}.$$

Крок 5. Перевіряємо умови зупинки алгоритму. Тобто, якщо  $\|U - U^*\|^2 < \epsilon$ , де  $U^*$  – матриця, отримана на попередньому кроці, то алгоритм закінчено, інакше переходимо до кроку 2.

У результаті роботи описаного алгоритму будуть знайдені міри належності об'єктів до кожного з кластерів, відповідно до заданих значень ознак об'єктів.

Таким чином, на основі отриманих даних можна буде зробити висновок про те, які об'єкти з початкової множини належать до ГВР.

Недоліком описаного підходу є те, що при проведенні такого поділу об'єктів на два кластери не враховуються властивості саме кластерів, які досліджуються.

Очевидно, що кінцевий результат значно залежить від початкового задання мір належності об'єктів до кожного з кластерів, і тому при різних початкових значеннях елементів матриці  $U^0 = (u_{ij})$  можливі неоднакові поділи об'єктів. Тому необхідною є розробка таких методів, які б дозволяли здійснювати виділення з сукупності ознак об'єктів тих ознак, що значною мірою впливають на належність до того чи іншого кластеру, та визначати правила для задання початкових мір належності об'єктів до кластерів. Також можливим є випадок, при якому початкова кількість кластерів буде більшою, ніж два, і після розбиття об'єктів на таку кількість кластерів відповідно до описаної вище схеми необхідно застосувати механізм об'єднання кластерів за деякими ознаками.

Такий підхід не є тривіальним та вимагає роботи як аналітиків в галузі інформаційних технологій, так і спеціалістів в галузі охорони здоров'я, соціології тощо.

## 5. Методи розв'язання задачі ідентифікації $Z_2$

Нейромережні технології. Розглянемо задачу  $Z_2$  оцінювання міри належності особи до ГВР.

Для даної задачі побудуємо нейронну мережу зустрічного поширення (НМЗП) так, як показано на рис. 2.

У даному випадку за допомогою нейронів шару Кохонена буде здійснюватися поділ об'єктів на два кластери відповідно до їх належності до ГВР.

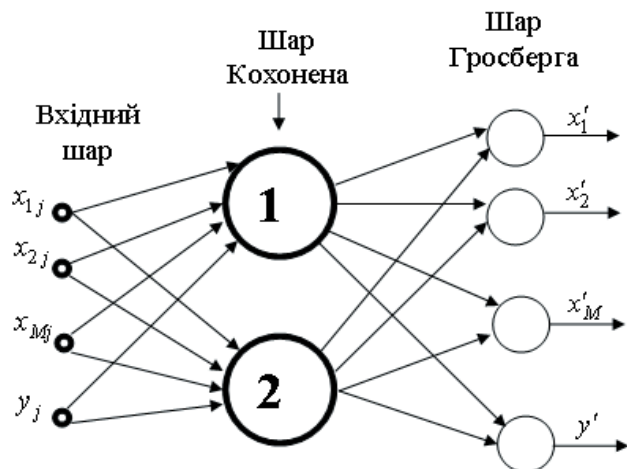


Рис. 2. Структурна схема НМЗП

Мережа описаного типу працює в двох режимах: навчання і використання [10]. В процесі навчання на входи необхідно подавати вектори характеристик об'єктів  $X^j = (x_{1j}, x_{2j}, \dots, x_{Mj})$ ,  $j = 1, N$  та початково задану множину  $Y$  – значень мір можливої належності кластеру. Аналогічно попередньому пункту проводиться коригування вагових коефіцієнтів нейронної мережі.

На етапі використання НМ на вхід подається вектор характеристик деякого об'єкта  $(x_1, x_2, \dots, x_N)$ , а на виході отримується  $y$  – міра належності даного об'єкта до досліджуваного кластеру.

Метод найменших квадратів. Розглянемо задачу ідентифікації  $Z_2$ . Припустимо, що за допомогою деякого методу для набору вхідних даних ми отримали множину  $Y = \{y_1, y_2, \dots, y_N\}$ , де  $y_j$  – міра належності кластеру об'єкта  $O_j$ . Для забезпечення можливості застосування МНК, необхідно провести дослідження мультиколінеарності даних, необхідно визначити, який вид регресії доцільно застосовувати – множинну лінійну регресію або ж множинну нелінійну регресію.

Основним критерієм “правильності” отриманої аналітичної залежності є мінімум суми квадратів відхилень табличних значень  $y_i$  від отриманих аналітично  $\tilde{y}_j$

$$E = \sum_{j=1}^N (y_j - \tilde{y}_j)^2 \rightarrow \min ,$$

де  $\tilde{y}_j = F(X^{(j)})$ .

Очевидно, що складність застосування даного підходу до розв'язання задачі ідентифікації міри належності особи до ГВР полягає у заданні початкових даних та обчисленні початкового вектору  $Y$ .

Недоліком застосування класичних методів ідентифікації для розв'язання задачі визначення міри належності особи до ГВР є те, що ще на перших етапах використання цих методів виникають значні труднощі при визначенні початкового вектору  $Y$  для множини осіб, що вже належать до ГВР.

Метод групового урахування аргументів (МГУА). Для застосування МГУА при розв'язуванні задачі

$Z_2$ , на попередньому етапі необхідно визначити множину  $Y = \{y_1, y_2, \dots, y_N\}$ , аналогічно до попереднього пункту.

З допомогою МГУА, за заданими вибірками здійснюється побудова великої кількості моделей, з яких, відповідно до введених критеріїв, і проводиться відбір моделі оптимальної складності.

В основі алгоритмів МГУА лежить індуктивний метод, згідно з яким дослідження ведеться від конкретних результатів спостереження до загальної моделі.

В загальному, даний метод – це метод побудови моделей на основі експериментальних даних в умовах невизначеності, причому побудова відбувається шляхом поділу вибірки на частини і перевірки адекватності проміжних моделей на різних підвбірках [17, 18].

Оптимальність моделі визначається критеріями селекції моделі, серед яких виділяють [17]: критерій регулярності; критерій незміщеності; критерій балансу змінних.

Виділяють наступні основні етапи МГУА:

1. Вибір опорної функції. Відповідно до апроксимаційної теореми Вейерштраса, будь-яку неперервну на відрізку функцію можна як завгодно точно наблизити поліномом з дійсними коефіцієнтами такого виду:

$$y = a_0 + \sum_{i=1}^M a_i x_i + \sum_{i=1}^M \sum_{j=1}^M a_{ij} x_i x_j + \sum_{i=1}^M \sum_{j=1}^M \sum_{k=1}^M a_{ijk} x_i x_j x_k + \dots$$

Як правило, опорною функцією обирають одну з таких:

$$y_k = a_0 + a_1 x_j ,$$

$$y_k = a_0 + a_1 x_i + a_2 x_i x_j ,$$

$$y_k = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j$$

тощо.

Крім того, є задачі [17], в яких за опорну функцію вибирають гармонічну, логічну функції, ряди Фур'є тощо.

2. Побудова та перебір моделей на основі різних опорних функцій. Моделі порівнюються згідно з критеріями для різних рядів чи груп даних і на основі такого порівняння обираються оптимальні за критеріями моделі.

3. Для декількох «кращих» відібраних моделей проводиться оцінка їх якості з використанням коефіцієнта детермінації, або коефіцієнта кореляції, або середньоквадратичного відхилення похибки. На основі проведеного аналізу відбирається потрібна модель і алгоритм закінчується, інакше – здійснюється перехід до наступного етапу.

4. Знайдені на попередніх етапах моделі розглядаються як аргументи опорних функцій і алгоритм починає нову ітерацію.

Недоліком застосування МГУА для задачі ідентифікації  $Z_2$  є те, що для оптимального вибору опорних функцій необхідні спеціальні знання та досвід аналітика.

---

## 6. Висновки

---

В роботі проаналізовано задачу оцінювання кількісних характеристик ГВР. Було проведено розбиття вказаної задачі на дві взаємопов'язані задачі: задачу ідентифікації і задачу кластеризації. Проаналізовано способи застосування деяких класичних методів ідентифікації та кластеризації для сформульованих задач.

В ході аналізу було показано, що застосування таких методів, як правило, супроводжується значними труднощами та не приводить до отримання вірних результатів. А, отже, необхідною є розробка нових моделей та методів для розв'язування задачі оцінювання кількісних характеристик ГВР, які б дозволяли враховувати такі джерела вхідних даних як професійна думка ОПР та висновки експертів.

---

## Література

1. Глобальный доклад: Доклад ЮНЭЙДС о глобальной эпидемии СПИДа [Текст]. – "ЮНЭЙДС", 2012. – 212 с.
2. Аналітичний звіт за результатами соціологічного дослідження "Оцінка чисельності груп високого ризику інфікування ВІЛ в Україні станом на 2009 рік" [Текст] / Г. Берлева, К. Думчев, Ю. Кобища та ін. – К.: МБФ "Міжнародний Альянс з ВІЛ/СНІД в Україні", 2010. – 104 с.
3. Моніторинг поведінки та поширеності ВІЛ серед осіб, які надають сексуальні послуги за плату, як компонент епідагляду за ВІЛ другого покоління: аналітичний звіт за результатами біоповедінкового дослідження 2011 року [Текст] / А. Грушецький – К.: МБФ «Міжнародний Альянс з ВІЛ/СНІД в Україні», 2012. – 120 с.
4. Моніторинг поведінки та поширеності ВІЛ серед жінок, які надають сексуальні послуги за плату, як компонент епідагляду за ВІЛ другого покоління: аналітичний звіт за результатами опитування ЖКС у 2009 році / А. Грушецький – К.: МБФ «Міжнародний Альянс з ВІЛ/СНІД в Україні», 2010. – 74 с.
5. Моніторинг поведінки жінок, які надають сексуальні послуги за плату як компонент епідагляду другого покоління [Текст] / О. Р. Артюх, О. М. Балакірева, Л. В. Бочкова та ін. – Київ: МБФ «Міжнародний Альянс з ВІЛ/СНІД в Україні», 2005. – 60 с.
6. Оцінка чисельності груп високого ризику інфікування ВІЛ в Україні станом на 2012 рік: аналітичний звіт за результатами дослідження [Текст] / Г. Берлева, К. Думчев, М. Касянчук та ін. – К.: МБФ «Міжнародний Альянс з ВІЛ/СНІД в Україні», 2012. – 68 с.
7. Оценка с привлечением к участию. Методические рекомендации [Текст] / под ред. О. М. Балакірева. – К.: МБФ «Международный Альянс по ВИЧ/СПИД в Украине», 2001. – 157с.
8. Мулеса, О. Ю. Нечітка ступінчата модель послідовного аналізу варіантів [Текст] / О. Ю. Мулеса // Вісник ЧДТУ. – 2012. – № 3. – С. 9-13.
9. Загоруйко, Н. Г. Методы распознавания и их применение [Текст] / Н. Г. Загоруйко. – М.: Изд-во "Советское радио", 1972. – 208 с.
10. Снитюк, В. Є. Прогнозування. Моделі. Методи. Алгоритми [Текст]: навчальний посібник / В. Є. Снитюк – К.: "Маклаут", 2008. – 364 с.
11. Ту, Дж. Принципы распознавания образов [Текст] / Дж. Ту, Р. Гонсалес. – М.: "Мир", 1978. – 412 с.
12. Гроп, Д. Методы идентификации систем [Текст] / Д. Гроп. – М.: Мир, 1979. – 302 с.
13. Kohonen, T. Self-organization and associative memory [Текст] / T. Kohonen. – New-York, Springer Verlag, 1989. – 312 p.
14. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика: учеб. пособие [Електронний ресурс] / Е. И. Большакова, Э. С. Клышинский, Д. В. Ландэ и др. – М.: МИЭМ, 2011. – 272 с. – Режим доступа: \www / URL: [http://www.webground.su/data/lit/bolshakova\\_klyshinsky\\_landé\\_noskov\\_peskova\\_yagunova/Avtomaticeskaya\\_obrabotka\\_tekstov.pdf](http://www.webground.su/data/lit/bolshakova_klyshinsky_landé_noskov_peskova_yagunova/Avtomaticeskaya_obrabotka_tekstov.pdf).
15. Штовба, С. Д. Введение в теорию нечетких множеств и нечеткую логику [Електронний ресурс] / С. Д. Штовба. – Режим доступа: \www/ URL: <http://matlab.exponenta.ru/fuzzylogic/book1/index.php>.
16. Смешко, Ю. В. Об одном критерии для выбора значения экспоненциального веса в алгоритме классификации нечетких С-средних [Електронний ресурс] / Ю. В. Смешко // Молодежь и наука. – Режим доступа: \www / URL: <http://conf.sfu-kras.ru/sites/mn2012/thesis/s012/s012-088.pdf>.
17. Ивахненко, А. Г. Индуктивный метод самоорганизации моделей сложных систем [Текст] / А. Г. Ивахненко. – К.: Наук. думка, 1981 – 296 с.
18. Степашко, В. С. Теоретические аспекты МГУА как методы индуктивного моделирования [Електронний ресурс] / В. С. Степашко. – Режим доступа: \www/ URL: <http://gmdh.net/articles/usim/Stepashko.pdf>.