

Олецький О. В., Тригуб О. С.

## ПРО ЗАСТОСУВАННЯ МЕТОДУ АНАЛІЗУ ІЄРАРХІЙ ДЛЯ АВТОМАТИЗОВАНОГО ОЦІНЮВАННЯ СТУДЕНТСЬКИХ РОБІТ

У статті досліджено задачу багатокритеріальної оптимізації, яку розв'язують методом аналізу ієрархій. Розглянуто ситуацію, коли потрібно порівняти та ранжувати студентські роботи (проекти), приблизно одного рівня якості. Продемонстровано, що в таких випадках класичний метод аналізу ієрархій, запропонований Т. Сааті, призводить до отримання досить «грубих» числових результатів (глобальних пріоритетів), коли виявляється, що робота одного студента суттєво переважає роботу іншого (хоча це не так, бо розглядали роботи приблизно одного рівня). Запропоновано використати альтернативну шкалу попарних порівнянь, яка даватиме змогу отримувати в таких випадках більш адекватні числові результати.

Проведено відповідні чисельні розрахунки із використанням авторської програмної системи, результати яких наведено у вигляді скріншотів.

**Ключові слова:** метод аналізу ієрархій, багатокритеріальна оптимізація, автоматизоване оцінювання, навчальний проект, попарні порівняння, транзитивні шкали, програмна система.

### Вступ

Нехай  $A = \{a_1, \dots, a_q\}$  – деяка множина альтернатив (варіантів), які потрібно порівнювати між собою, і, можливо, вибирати з цих альтернатив найкращу. Постановку задачі оцінювання та ранжування альтернатив можна вважати задовільно формалізованою, якщо на множині альтернатив можна адекватним способом задати таку функцію  $u: A \rightarrow R$ , що варіант  $a$  краще за варіант  $b$  тоді і тільки тоді, коли  $u(a) > u(b)$ . Подібний підхід є характерним для сучасних напрямів, пов'язаних із прийняттям рішень в інтелектуальних системах; функцію  $u$  при цьому часто називають функцією корисності [9]. Але таку функцію вдається задати далеко не завжди.

Дуже типовою є спроба представлення функції  $u$  у вигляді зваженої суми (згортки) критеріїв (ознак)

$$u(a_i) = \sum_{j=1}^n \lambda_j a_{ij}, \quad (1)$$

де  $a_{ij}$  – значення  $j$ -ї ознаки для  $i$ -ї альтернативи,  $\lambda_j$  – вагові коефіцієнти, які задають міри важливості ознак. Однак раціональний вибір ознак та мір їх важливості часто також є проблемою.

Подібні ситуації нерідко виникають при оцінюванні успішності студентів. Особливо це стосується завдань, які мають неформальний характер (наприклад, навчальні проекти, які передбачають створення програмних засобів із широкою функціональністю, доволі дружнім користувацьким інтерфейсом тощо). Функціональність, друж-

ність інтерфейсу, вибір засобів реалізації часто оцінюється інтуїтивно на рівні «подобається – не подобається». Далі постає закономірне запитання: викладач може вважати таку інтуїтивну оцінку адекватною, але її потрібно обґрунтувати, інакше можуть виникати конфліктні ситуації. І навіть якщо вдається встановити ті чи ті формальні критерії, залишається серйозне питання про адекватну оцінку мір їхньої важливості.

Водночас дедалі актуальнішим стає питання про автоматизоване оцінювання виконання завдань студентами. Таке оцінювання могло б здійснюватися з використанням певних алгоритмів на основі експертних оцінок, виставлених викладачами та запрошеними спеціалістами. При цьому треба враховувати, що експертам, як правило, легше порівняти два варіанти між собою, ніж давати якісь кількісні оцінки.

Одним із найбільш рекомендованих методів, який добре зарекомендував себе для вирішення подібних проблем, є метод аналізу ієрархій, запропонований Т. Сааті. У своїй найпростішій формі, орієнтованій на оцінювання варіантів за формулою (1), цей метод можна охарактеризувати таким способом [5; 9].

1. Для довільного набору елементів, для якого можна задати відношення переваг, вектор чисельних оцінок елементів можна отримати як нормалізований головний власний вектор квадратної обернено-симетричної матриці попарних порівнянь  $M$ , де  $m_{ij}$  – міра переваги  $i$ -го елемента над  $j$ -м відповідно до певної шкали градацій переваг.

2. Значення  $a_{ij}$  отримуються на основі попарних порівнянь альтернатив окремо за кожним критерієм.

3. Коефіцієнти  $\lambda_i$  отримуються на основі попарних порівнянь самих критеріїв.

Більш формально це можна записати так.

Варіанти (елементи множини  $A$ ) оцінюються на основі попарних порівнянь за кожним із  $n$  критеріїв  $f_j(a)$ . В результаті отримуються коефіцієнти  $a_{ij} = f_j(a_i)$ .

Далі на основі попарних порівнянь обчислюються ваги самих критеріїв; отримується вектор  $\lambda = (\lambda_1, \dots, \lambda_n)$ .

Нарешті, оптимальний варіант отримується як

$$a^* = \arg \max_i J(a_i),$$

де

$$J(a_i) = \sum_{j=1}^n \lambda_j f_j(a_i) = \sum_{j=1}^n \lambda_j a_{ij}.$$

Дуже важливого значення при цьому набуває вибір чисельних значень (градацій) мір переваг, які використовують для формування матриці  $M$ . Т. Сааті запропонував таку шкалу градацій:

1 – варіанти рівнозначні;

3 – слабка перевага;

5 – сильна перевага;

7 – дуже сильна перевага;

9 – абсолютна перевага.

Значення 2, 4, 6, 8 розглядаються як проміжні.

Однак механічне застосування такої шкали не завжди дає змогу отримати добрі результати. На це звертали увагу багато дослідників [9], і ситуація в цьому питанні досить суттєво залежить від конкретної предметної сфери.

Зовсім не обов'язково постулювати, що якщо варіант 1 має слабку перевагу над варіантом 2, а варіант 2 має слабку перевагу над варіантом 3,

то 1 повинен мати сильну перевагу над варіантом 3, – ця перевага цілком може бути оцінена як слабка. Крім того, відношення переваг взагалі може не бути транзитивним.

### Постановка задачі та основний зміст роботи

Розглянемо таку ситуацію. Нехай потрібно на основі методу аналізу ієрархії оцінити чотири студентські проекти. Кожний проект оцінюється за трьома критеріями, які експерти оцінюють як однаково важливі.

За кожним із критеріїв проекти лінійно впорядковані, тобто

$$1 \succ 2 \succ 3 \succ 4.$$

Однак ці переваги є незначними. Більше того, всі проекти є достатньо якісними й заслуговують на високу оцінку.

Відповідно до стандартної шкали Т. Сааті, ми маємо оцінити мінімальну градацію переваги значенням 2. Оскільки всі переваги мінімальні, обернено-симетричні матриці переваг за кожним із трьох критеріїв матимуть вигляд:

$$\begin{pmatrix} 1 & 2 & 2 & 2 \\ \frac{1}{2} & 1 & 2 & 2 \\ \frac{1}{2} & \frac{1}{2} & 1 & 2 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}$$

Для проведення чисельного моделювання використаємо авторську програмну систему підтримки прийняття управлінських рішень «SoftExpert v.2.0» [6–8; 10]. Після введення кількості критеріїв та альтернатив, система побудує представлення задачі у вигляді ієрархії (рис. 1).



Рис. 1. Представлення модельної задачі у вигляді ієрархії

Попарні порівняння елементів задачі

**Попарні порівняння відносно "КОНКУРС"**

	1	2	3	Локальні пріоритети
1. КРИТЕРІЙ 1	1	1	1	0.33
2. КРИТЕРІЙ 2	1	1	1	0.33
3. КРИТЕРІЙ 3	1	1	1	0.33

Рис. 2. Заповнена матриця попарних порівнянь критеріїв

Усі три критерії мають рівну вагомість, тому за їх попарного порівняння матимемо квадратну матрицю розмірності 3×3, заповнену одиницями (рис. 2).

Аналогічно введемо до системи і три матриці попарних порівнянь альтернатив (за кожним критерієм – своя матриця) (рис. 3).

Як бачимо у правому верхньому стовпчику на рис. 3, за такого задання переваг система вирахувала нормалізований вектор мір важливостей (вектор локальних пріоритетів), який має вигляд (із точністю до другого знаку після коми):

0,39 0,28 0,20 0,14.

При натисканні кнопки «Розрахувати» (див. рис. 1) система виводить результат розрахунків (у вигляді гістограми) (рис. 4).

Нехай максимально можлива оцінка за проєкт дорівнює 30 балам (така ситуація є дуже типовою). Якщо прийняти, що найкращий проєкт заслуговує на найвищу оцінку, а інші оцінки виставляють пропорційно, маємо такий розподіл (з округленням до цілих чисел):

30 21 15 11.

Таке виставлення оцінок не відповідає інтуїтивному уявленню про «природність» оцінок, оскільки різниця між максимальним і мінімальним

Попарні порівняння елементів задачі

**Попарні порівняння відносно "КРИТЕРІЙ 1"**

	1	2	3	4	Локальні пріоритети
1. СТУДЕНТ 1	1	2	2	2	0.39
2. СТУДЕНТ 2	1/2	1	2	2	0.28
3. СТУДЕНТ 3	1/2	1/2	1	2	0.20
4. СТУДЕНТ 4	1/2	1/2	1/2	1	0.14

Порівнюємо "СТУДЕНТ 3" з "СТУДЕНТ 4"

Елемент переважає	Елемент поступається
Рівна вагомість (1)	Рівна вагомість (1)
Проміжний рівень (2)	Проміжний рівень (1/2)
Переважає помірно (3)	Постається помірно (1/3)
Проміжний рівень (4)	Проміжний рівень (1/4)
Переважає суттєво (5)	Постається суттєво (1/5)
Проміжний рівень (6)	Проміжний рівень (1/6)
Переважає сильно (7)	Постається сильно (1/7)
Проміжний рівень (8)	Проміжний рівень (1/8)
Переважає абсолютно (9)	Постається абсолютно (1/9)

Величина узгодженості даних складає 4.49 %

Рис. 3. Заповнена матриця попарних порівнянь альтернатив (на прикладі критерію 1)

значенням є надто великою, і проєкт, який посів останнє місце, отримує незаслужено низький бал. Якщо ж у матрицях попарних порівнянь оцінити всі варіанти як рівнозначні за кожним із критеріїв (заповнити комірки матриць значенням 1), то оцінки будуть однаковими, що також неправильно. Змагальну компоненту при оцінюванні студентських проєктів має бути збережено, і попарні порівняння – хороший шлях до цього.

Можливий вихід із ситуації міг би полягати в застосуванні до отриманого вектора оцінок деякого нелінійного згладжувального перетворення, яке б зменшило розкид між максимальним та мінімальним значенням. Але можна піти іншим шляхом.

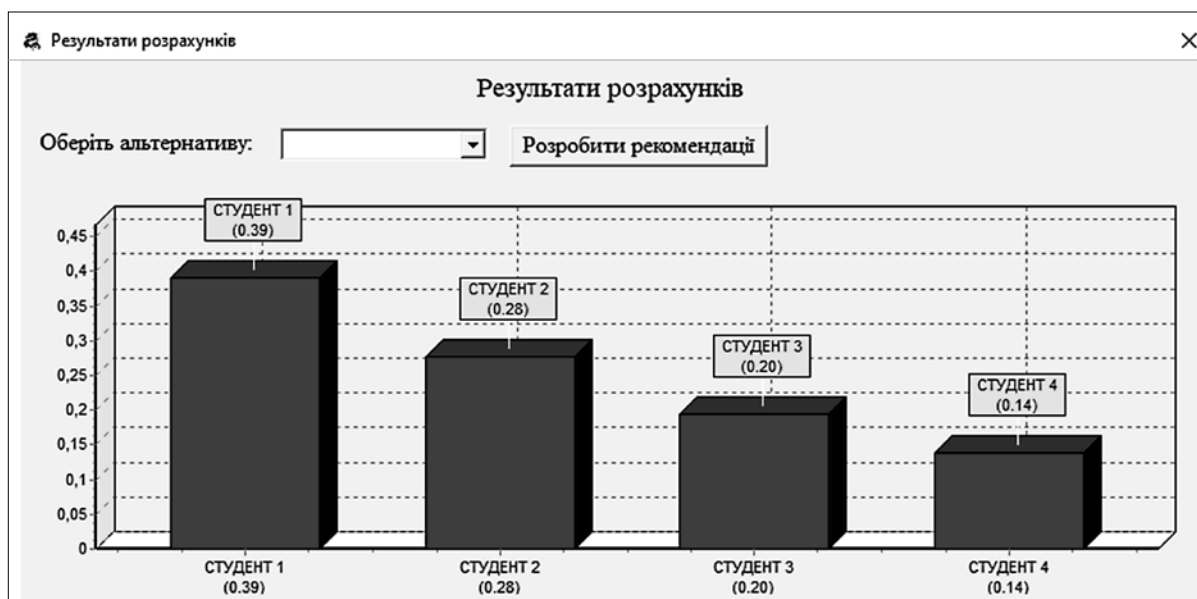


Рис. 4. Результати розрахунків

Ідея полягає у використанні альтернативної шкали переваг, як це було описано в [9]. Вводиться деякий параметр  $\tau > 1$ . Тоді рівноцінність варіантів у матриці попарних порівнянь буде оцінено значенням  $1$ , мінімальна перевага – значенням  $\tau$ , наступна градація переваг – значенням  $\tau^2$  і т. д. Іншими словами, можна розглядати такі градації переваг:

- $1 = \tau^0$  – альтернативи рівноцінні, немає переваги;
- $\tau^1$  – слабка перевага;
- $\tau^2$  – сильна перевага;
- $\tau^3$  – дуже сильна перевага;
- $\tau^4$  – абсолютна перевага.

Подібні шкали названі в [9] транзитивними. Очевидно, що для кожної конкретної задачі кількість таких градацій може бути змінено.

Розглянемо можливість застосування такої шкали до задачі автоматизованого оцінювання студентських робіт. У наведеному прикладі матриці переваг за окремими критеріями матимуть вигляд:

$$\begin{pmatrix} 1 & \tau & \tau & \tau \\ \frac{1}{\tau} & 1 & \tau & \tau \\ \frac{1}{\tau} & \frac{1}{\tau} & 1 & \tau \\ \frac{1}{\tau} & \frac{1}{\tau} & \frac{1}{\tau} & 1 \end{pmatrix}$$

При  $\tau = 1.1$  розподіл оцінок з округленням до цілих має вигляд:

$$30 \quad 29 \quad 27 \quad 26,$$

що більш-менш відповідає інтуїтивним уявленням про «природність» оцінок.

### Висновки та обговорення

Вибір параметра  $\tau$  є окремим і недостатньо дослідженим питанням. Може бути поставлено завдання підібрати цей параметр так, щоб оцінки розставлялися більш-менш «природним» чином, і таке оцінювання не викликало б серйозних обґрунтованих заперечень у викладачів, експертів та самих студентів. Як можливий підхід до формалізації цього критерію можна запропонувати такий підбір  $\tau$ , щоб різниця між максимальною та мінімальною оцінкою потрапляла в деякий діапазон середніх значень: була не надто великою, але й не надто маленькою. У порівняно простих випадках «найкраще» значення  $\tau$  може бути знайдено на основі простого перебору або половинного ділення. Для такої формалізації, своєю чергою, може бути запропонований підхід на основі застосування нечітких множин [1; 2; 4]. Це питання має стати предметом подальшого дослідження.

### Список літератури

1. Глибовець М. М. Штучний інтелект / М. М. Глибовець, О. В. Олецький. – Київ: Вид. дім «Академія», 2002. – 366 с.
2. Олецький О. В. Про підхід до автоматичного формування рекомендацій для відвідувачів веб-порталу на основі теорії нечітких множин / О. В. Олецький // Наукові записки НаУКМА. Комп'ютерні науки. – 2015. – Т. 177. – С. 37–40.
3. Рассел С. Искусственный интеллект: современный подход / С. Рассел, П. Норvig. – Москва: Изд. дом «Вильямс», 2006. – 1408 с.
4. Рутковская Д. Нейронные сети, генетические алгоритмы и нечеткая логика / Д. Рутковская, М. Пилинский, Л. Рутковский. – Москва: Горячая линия - Телеком, 2004. – 452 с.
5. Саати Т. Принятие решений. Метод анализа иерархий / Т. Саати. – Москва: Радио и связь, 1993.
6. Тригуб О. С. Програмна система дослідження слабоструктурованих задач багатокритеріальної оптимізації / О. С. Тригуб // XII Міжнародна науково-практична конференція «Теоретичні та прикладні аспекти побудови програмних систем» – ТАAPSD'2015: тези доп., 23–26 листопада 2015 р. – Київ, 2015. – С. 176–180.
7. Тригуб О. С. Система аналізу та прийняття управлінських рішень SoftExpert / О. С. Тригуб // XIV Міжнародна науково-практична конференція «Теоретичні та прикладні аспекти побудови програмних систем» – ТАAPSD'2017: 4–8 грудня 2017 р., Київ. – Київ, 2017. – С. 178–179.
8. Тригуб О. С. Використання та програмна реалізація методу аналізу ієрархій при розв'язанні багатокритеріальних задач / О. С. Тригуб, Р. О. Тригуб // Міжнародна наукова конференція «Сучасні проблеми математичного моделювання, обчислювальних методів та інформаційних технологій» (присвячена пам'яті акад. І. І. Ляшка): 2–4 березня 2018 р., Рівне. – Рівне, 2018. – С. 90–92.
9. Черноруцкий И. Г. Методы принятия решений / И. Г. Черноруцкий. – Санкт-Петербург: БХВ-Петербург, 2005. – 416 с.
10. Tryhub O. S. Researching semistructured problems of multicriteria optimization using the software system / O. S. Tryhub, R. O. Tryhub, V. Gorbokov // Наукові записки НаУКМА. Комп'ютерні науки. – 2013. – Т. 151. – С. 79–88.

### References

- Chernorutskiy, Y. (2005). *Metody prinyatiia resheniy*. Sankt-Peterburg: BKhV-Peterburh [in Russian].
- Hlybovets, M., & Oletskiy, O. (2002). *Shtuchnyi intelekt*. Kiyv: Vyd. dim "Akademiiia" [in Ukrainian].
- Oletskiy, O. (2015). Pro pidkhdid do avtomatychnoho formuvannia rekomendatsii dlia vidviduvachiv veb-portaluu na osnovi teorii nechitkykh mnozhyn. *Naukovi zapysky NaUKMA. Kompiuterni nauky*, 177, 37–40 [in Ukrainian].
- Rassel, S., & Norvyh, P. (2006). *Iskusstvennyi intellekt: sovremennyi podkhod*. Moskva: Izd. dom "Vil'jams" [in Russian].
- Rutkovskaia, D., Pilinskiy, M., & Rutkovskiy, L. (2004). *Neironnye seti, heneticheskie alhoritmy v nechetkoi lohyke*. Moskva: Horiachaia liniia – Telekom [in Russian].
- Saaty, T. (1993). *Pryniatyie resheniy. Metod analiza yerarkhyyi*. Moskva: Radio i sviaz [in Russian].
- Tryhub, O. (2015). Prohramna sistema doslidzhennia slabostrukturnykh zadach bahatokryterialnoi optymizatsii. In *XII Mizhnarodna naukovo-praktychna konferentsiia "Teoretychni ta prykladni aspekty pobudovy prohramnykh system"* – TAAPSD2015: tezy dop., 23–26 lystopada 2015 r. (pp. 176–180). Kyiv [in Ukrainian].

Tryhub, O. (2017). Systema analizu ta pryiniattia upravlinskykh rishen SoftExpert. In *XIV Mizhnarodna naukovo-praktychna konferentsiia "Teoretychni ta prykladni aspekty pobudovy proqramnykh system" – TAAPSD'2017: 4–8 hrudnia 2017 r.*, Kyiv (pp. 178–179). Kyiv [in Ukrainian].

Tryhub, O. S., Tryhub, R. O., & Gorburokov, V. (2013). Researching semistructured problems of multicriteria optimization using

the software system. *Naukovi zapysky NaUKMA, Kompiuterni nauky*, 151, 79–88 [in Ukrainian].

Tryhub, O., & Tryhub, R. (2018). Vykorystannia ta prohramna realizatsiia metodu analizu iierarkhii pry rozviazanni bahatokryterialnykh zadach. In *Mizhnarodna naukova konferentsiia "Suchasni problemy matematychnoho modeliuвання, obchysliuvalnykh metodiv ta informatsiinykh tekhnolohii" (prysviachena pamiaty akad. I. I. Liashka): 2–4 bereznia 2018 r.*, Rivne (pp. 90–92). Rivne [in Ukrainian].

O. Oletsky, O. Tryhub

## USE OF ANALYTICAL HIERARCHY PROCESS FOR AUTOMATED ASSESSING OF STUDENTS' STUDY PROJECTS

*The problem of applying multicriteria optimization on the basis of the analytical hierarchy process (AHP) for a specific subject domain is regarded in the paper. Namely, the situation when some students' study projects should be compared and assessed with the aid of automated algorithm-based tools for grading especially if these projects are nearly of the same quality level is considered. It is well-known that using the classical graduation scale of pairwise comparisons preferences, suggested by T.Saati and commonly used in AHP, may lead to not very good results. The reason for this is that the difference between the automatically evaluated grades of the projects may become unacceptably large even though all of the projects are good enough and deserve good grades. On the other hand, a competitive and comparative approach to grading study projects should be preserved.*

*In order to mitigate this problem, many authors proposed to apply alternative scales of pair comparisons, which are sometimes referred to as transitive ones. Constructing such scales should imply introducing some parameter that determines how many times the next graduation in the scale of preferences in pair comparisons matrices is bigger than the previous one. Typically this parameter is a rational value between 1 and 2 whereas 2 is the least graduation of preferences in the standard Saati scale. The main issue is how to find out appropriate values of this parameter in order to meet the requirements of the certain task.*

*For exploring this question, some numerical experiments have been carried out and are reported in this paper. According to the typical algorithm of estimations, some projects executed by students have been compared pairwise by each criterion separately. Then the chosen criteria have been compared pairwise as well, and the linear convolution method has been applied for getting final grades. The reported experiments show that for the task of automated grading of students' study projects alternative transitive scales described above can provide more adequate results than using the standard Saati scale, and the best choice of the parameter for this specific task is about 1.1.*

*All experiments have been carried out with the help of the author software toolkits developed for the purposes of exploring AHP and of applying it for various practical tasks. Some screenshots illustrating the main points of reported experiments are supplied.*

**Keywords:** analytical hierarchy process, multicriteria optimization, automated assessing, study project, pair comparisons, transitive scales, software system.



Creative Commons Attribution 4.0 International License (CC BY 4.0)

Матеріал надійшов 20.05.2020