

УДК 519.6 : 519.233.6

DOI: 10.15587/2313-8416.2015.51512

## КОМПЛЕКСНИЙ АНАЛІЗ СОЦІАЛЬНО-ЕКОНОМІЧНОГО СТАНУ РЕГІОНІВ З ВИКОРИСТАННЯМ GRID-ТЕХНОЛОГІЙ

©Л. В. Зубик, О. М. Гладка, І. М. Карпович, В. О. Савич

*На основі інтелектуального аналізу даних з використанням Grid-технологій проведено ранжування соціально-економічного стану районів Рівненської області за статистичними показниками їх господарської діяльності у першому півріччі 2015 року. Виконано системне оцінювання із застосуванням елементів методів шкалювання, факторного та кластерного аналізу. Спроектовано гнучкий і масштабований модуль, який розширює можливості стандартних web-сервісів і спрощує процедури збору та опрацювання даних*

**Ключові слова:** *аналіз даних, ранжування, інтелектуальна система оцінювання, багатовимірне шкалювання, Grid-технологія*

*Ranking of the socioeconomic conditions of districts of Rivne region on statistical indicators of economic activity in the first half of 2015 was held based on Data Mining using Grid-technologies. System evaluation using the elements of methods of scaling, factor and cluster analysis was performed. Flexible and scalable module that extends the capabilities of the standard web-services and simplifies the procedures of collection and processing of data was designed*

**Keywords:** *data analysis, ranking, intelligent evaluation, multidimensional scaling, Grid-technology*

### 1. Вступ

Сучасний етап розвитку інформаційного суспільства в умовах глобалізації характеризується активним застосуванням досягнень інформатики та обчислювальної техніки для розв'язування актуальних соціально-економічних і екологічних проблем з метою забезпечення сталого розвитку і безпеки. На реалізацію таких задач як в регіональному, так і в глобальному масштабі спрямовано багато міжнародних програм і ініціатив. Зокрема в Україні, як і у всьому світі, отримують розвиток розподілені інформаційні системи, які надають доступ до великих сховищ інформації та інших ресурсів, а також передбачають сумісне використання інформаційних продуктів і інтеграцію регіональних та національних систем моніторингу [1, 2].

### 2. Аналіз літературних даних та постановка проблеми

Розвиток ринкової економіки супроводжується корінними змінами структури і зв'язків народного господарства, нестабільністю, загостренням певних негативних явищ. Адекватна реакція на ці явища з боку органів державного управління можлива лише за наявності системи об'єктивного відбиття і оперативної передачі досить повної мікро- і макрорівневої інформації. З іншого боку, виникнення різноманітних форм власності, поява багатьох нових підприємств, в

т. ч. малих, висуває підвищені вимоги до достовірності і оперативності інформації. Приватному підприємцю необхідна об'єктивна картина розвитку бізнесу і обґрунтований перспективний прогноз. Зокрема, для розміщення інвестицій необхідне знання не тільки економічного стану, але і реальної оцінки рівнів соціальної, політичної стабільності, економічності промислових об'єктів тощо.

Для відображення і моделювання реальних явищ і процесів в аналітичній практиці останніх років активно використовується Grid-технологія [3, 4]. Прогнозування розвитку і підтримка прийняття відповідних рішень забезпечується тут ефективним використанням, опрацюванням і аналізом даних з різних джерел, а об'єкти, соціальні і економічні явища розглядаються із врахуванням деякої множини ознак. Використання методів багатовимірного статистичного аналізу [5, 6] дозволяє добиватися повноти теоретичного опису об'єктів дослідження і об'єктивності наступних висновків.

Grid-система, як відомо, – це програмно-апаратне середовище, побудоване на основі обчислювальних пристроїв, що належать різним адміністративним доменам комунікаційної мережі, яка дозволяє дистанційно використовувати інформаційні і обчислювальні ресурси цих пристроїв (обсяг пам'яті, можливості процесора, програмне забезпечення, бази даних тощо). Такі системи створені і успішно вико-

ристовуються в різних прикладних галузях. Дослідження соціально-економічних систем висуває специфічні вимоги до побудови і функціонування Grid-систем. Тому увага авторів цієї статті була зосереджена на розвитку з позицій системного аналізу методики побудови Grid-систем із врахуванням особливостей прикладних задач дослідження соціально-економічних систем.

### 3. Ціль та задачі дослідження

Метою даної роботи є аналіз економічного стану північних районів Рівненської області за розрахованим комплексним показником рівня розвитку їх економічної діяльності. В основу розрахунків покладено статистичні показники районів, а саме – виробництво основної сільськогосподарської продукції господарствами всіх категорій та фінансовий

стан господарств за підсумками першого півріччя 2015 року [7].

### 4. Матеріали та методи дослідження

Вибір об'єктів дослідження зумовлений, головним чином, подібністю кліматичних, природних та інших умов господарювання. На нашу думку, для аналізу відібрано основні фактори економічної діяльності господарств (табл. 1).

При проведенні порівняльного аналізу (рангові оцінки) використовують таксономічні методи, методи шкалування, деякі методи факторного та кластерного аналізу [8–10]. Інформаційні технології і нові досягнення в галузі високопродуктивних обчислень, які доступні тепер зацікавленим особам практично у всіх сферах людської діяльності, значно підвищили ефективність і якість розв'язування таких задач.

Таблиця 1

Економічне і соціальне становище Рівненської області за січень-липень 2015 р.

Райони Рівненської області	Обсяг реалізованої промислової продукції (товарів, послуг) без ПДВ та акцизу за січень-червень 2015 р., тис. грн.	Урожайність зернових та зернобоб. культур у с/г підприємствах на 01.08.15 р., ц з 1 га зібр. площі	Виробництво м'яса у с/г підприємствах на 01.08. 2015 р., т	Виробництво молока у с/г підприємствах на 01.08. 2015 р., т	Прийнято в експлуатацію житла по містах та районах за січень-червень 2015 р., кв. м	Прямі іноземні інвестиції (акціонерний капітал) у містах та районах області на 01.07.15 р., тис. дол. США	Заборгованість населення з платежів за комунальні послуги на 01.07. 2015 р., тис. грн.	Кількість зареєстрованих безробітних станом на 01.07. 2015 р., осіб
Березнівський	157090,3	22,4	3041	331	7795	824,4	5811,9	767
Володимирецький	22698,9	14,3	28	289	8337	2022,8	256,3	734
Дубровицький	30313,0	31,7	110	91,1	3199	59,2	1493,0	505
Зарічненський	3098,7	21,0	24	69,5	500	50,0	220,6	633
Костопільський	697420,9	19,1	81	130,2	8246	4466,8	2518,1	851
Рокитнівський	189105,6	14,5	22	78	4414	1320,7	2455,1	1105
Сарненський	277658,2	19,2	117	96,3	13317	16609,7	5557,3	1216

Для розв'язування поставленої задачі сформулюємо алгоритм обчислення таксономічного показника рівня розвитку багатовимірних об'єктів, який дозволяє впорядковувати (ранжувати) елементи даної сукупності.

Нехай дано множину із  $m$  об'єктів, кожен з яких описується  $n$  ознаками. Тоді кожен об'єкт дослідження можна розглядати як точку  $n$ -мірного простору, координатами якої є величини ознак. Координати точок складають матрицю спостережень:

$$X = \begin{bmatrix} x_{11} & \dots & x_{1n} \\ \dots & & \dots \\ x_{m1} & \dots & x_{mn} \end{bmatrix}, \quad (1)$$

де  $x_{ij}$  – значення ознаки  $j$  для  $i$ -го об'єкта.

1. Перейдемо від матриці спостережень до матриці стандартизованих даних за формулою:

$$2. z_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_j},$$

$$\text{де } \bar{x}_j = \frac{1}{m} \sum_{i=1}^m x_{ij}, \quad \sigma_j = \left[ \frac{1}{m} \sum_{i=1}^m (x_{ij} - \bar{x}_j)^2 \right]^{\frac{1}{2}},$$

тут  $\bar{x}_j$  – середнє арифметичне значення ознаки  $j$ ;

$\sigma_j$  – стандартне відхилення ознаки  $j$ .

$$Z = \begin{bmatrix} z_{11} & \dots & z_{1n} \\ \dots & & \dots \\ z_{m1} & \dots & z_{mn} \end{bmatrix}. \quad (2)$$

3. Розрахуємо коефіцієнти ієрархії. Для цього:

3. 1. Побудуємо матрицю відстаней  $D$  за формулою евклідової метрики:

$$d_{rs} = \left[ \sum_{k=1}^n (z_{rk} - z_{sk})^2 \right]^{\frac{1}{2}},$$

де  $d_{rs}$  – відстань між елементами  $r$  та  $s$ ;  $r=1..m$ ;  $s=1..m$ .

$$D = \begin{bmatrix} 0 & d_{12} & \dots & d_{1m} \\ d_{21} & 0 & \dots & d_{2m} \\ \dots & \dots & \dots & \dots \\ d_{m1} & d_{m2} & \dots & 0 \end{bmatrix}. \quad (3)$$

3. 2. Визначимо критичну відстань  $k$ . Для цього знаходимо найменшу ненульову відстань у кожному рядку або стовпці матриці  $D$  і вибираємо з них найбільше значення:

$$k = \max_r \min_s d_{rs}. \quad (4)$$

3. 3. Складемо множину  $F_r$  із тих елементів матриці  $D$ , які не перевищують критичну відстань  $k$ :

$$F_r = \{(r, s) | d_{rs} \leq k; s = 1..m\}. \quad (5)$$

3. 4. Далі обчислюємо:

$$g_r = \sum_{(r,s) \in F_r} d_{rs}; \quad g_{max} = \max_r g_r; \quad \lambda_r = \frac{g_r}{g_{max}} (r = 1..m).$$

4. Одержані значення коефіцієнтів ієрархії  $\lambda_r$  домножимо на величини відповідних ознак у стандартизованій матриці (2). Дістанемо матрицю

$$5. Z' = \{z'_{ij} | i = 1..m; j = 1..n\}.$$

6. Проведемо диференціацію ознак матриці спостережень. Ознаки, що позитивно впливають на рівень розвитку об'єктів, називають стимуляторами; ознаки, які перешкоджають розвитку об'єктів, називають дестимуляторами. Зокрема, в табл. 1 стимулятори економічної діяльності районів розміщені в стовпцях 2–7, а дестимулятори – в стовпцях 8–9.

7. На основі п. 4 побудуємо еталон розвитку, який являє собою точку  $E_0(z_{01}, z_{02}, \dots, z_{0n})$ , де  $z_{0s} = \max_r z'_{rs}$ , якщо  $s \in Y$ ;  $z_{0s} = \min_r z'_{rs}$ , якщо  $s \notin Y$ ; множина стимуляторів  $z'_{rs}$  – домножене на стандартизоване значення ознаки  $s$  для об'єкта  $r$ .

8. Розраховуємо відстань між окремими точками-об'єктами і точкою  $E_0$  за формулою:

$$b_{i0} = \left[ \sum_{j=1}^n (z'_{ij} - z_{0j})^2 \right]^{1/2}, \quad i = 1..m. \quad (6)$$

9. Знайдені відстані дозволяють розрахувати показник рівня розвитку:

$$10. \quad p_i = 1 - \frac{b_{i0}}{b_0}, \quad (7)$$

де

$$b_0 = \bar{b}_0 + 2\sigma_0; \quad \bar{b}_0 = \frac{1}{m} \sum_{i=1}^m b_{i0}; \quad \sigma_0 = \left[ \frac{1}{m} \sum_{i=1}^m (b_{i0} - \bar{b}_0)^2 \right]^{1/2}.$$

Зміст показника рівня розвитку  $p_i$  в наступному:  $i$ -й об'єкт тим більше розвинутий, чим ближче значення показника  $p_i$  до одиниці.

Таким чином, за допомогою обчисленого комплексного показника рівня розвитку можна оцінити відносний рівень об'єкта чи явища, що досліджується.

## 5. Результати досліджень

За розробленим алгоритмом складена і реалізована програма алгоритмічною мовою Visual Basic.

В табл. 2 наведені результати виконання цієї програми за вихідними даними табл. 1. Розраховані оцінки  $p_i$  дають можливість виокремити три категорії районів за їх економічним станом: I – достатньо розвинуті; II – помірно розвинуті; III – слабо розвинуті.

Зрозуміло, що кількість об'єктів дослідження і число показників кожного об'єкту може бути довільною. Програмою передбачена можливість проведення числового експерименту, тобто дослідження, як збільшення чи зменшення того чи іншого показника впливає на рейтинг об'єктів.

Таблиця 2

Ранжування районів Рівненської області за показником економічного стану

№	Район	Показник економічного стану ( $p_i$ )	Категорія стану
1	Костопільський	0,42	I
2	Березнівський	0,34	
3	Дубровицький	0,26	II
4	Сарненський	0,26	
5	Володимирецький	0,20	III
6	Зарічненський	0,14	
7	Рокитнівський	0,03	

## 6. Висновки

Виконано комплексну порівняльну оцінку соціально-економічного стану окремих регіонів Рівненської області, подібних за кліматичними, природними та іншими умовами господарювання, на основі показників їх діяльності протягом першого півріччя 2015 року, з позицій системного аналізу з використанням методів статистичного опрацювання даних. Для реалізації поставленого завдання розроблено методику системного оцінювання, яка базується на застосуванні комбінації наступних методів статистичного опрацювання даних: метод шкалування, методи факторного та кластерного аналізу. Спроектовано гнучкий і масштабований модуль для реалізації запропонованої методики оцінювання і розширення можливостей стандартних web-сервісів за рахунок використання переваг Grid-технологій. Запропонована методика ранжування об'єктів та практично реалізоване застосування спрощують процедури збору і аналізу даних для користувачів.

Комплексна оцінка економічного стану регіонів може бути використана як управлінськими структурами різних рівнів, так і суб'єктами господарювання для аналізу і прогнозування бюджетної, інвестиційної та іншої діяльності.

## Література

1. Цільова комплексна програма наукових досліджень НАН України “Грид-інфраструктура і грид-технології для наукових і науково-прикладних застосувань” [Електрон-ний ресурс]. – Постанова Президії НАН

України від 11.12.2013. – 2013. – № 164-а. – Режим доступу: <http://www1.nas.gov.ua/infrastructures/Legaltexts/nas/2013/Pages/>

2. Аналіз сталого розвитку – глобальний і регіональний контексти [Текст]: монографія / за ред. М. З. Згуровський, В. Я. Шевчук. – Міжнародна рада з науки (ICSU), Світовий центр даних з геоінформатики та сталого розвитку, Інститут прикладного системного аналізу НАН України і МОН України. – Глобальний аналіз якості та безпеки життя людей. – К.: НТУУ “КПІ”, 2010. – 252 с.

3. Куссуль, Н. М. Grid-системи для задач дослідження Землі. Архітектура, моделі і технології [Текст] / Н. М. Куссуль, А. Ю. Шелестов. – К.: Наукова думка, 2008. – 452 с.

4. Згуровський, М. З. Створення національної Grid-інфраструктури для забезпечення наукових досліджень [Текст] / М. З. Згуровський, А. І. Петренко, Г. Д. Кисельов // Інформаційні технології в освіті. – 2009. – № 4. – С. 12–17.

5. Сошникова, Л. А. Многомерный статистический анализ в экономике. [Текст] / Л. А. Сошникова, В. Н. Тамашевич, Г. Уебе, М. Шефер. – М.: ЮНИТИ – ДАНА, 1999. – 598 с.

6. Backhaus, K. Multivariate Analysemethoden: Eine anwendungsorientierte Einführung [Text] / K. Backhaus, B. Erichson, W. Plinke, R. Weiber. – Springer Berlin, 2006. – 830 p.

7. Економічне і соціальне становище Рівненської області за січень-липень 2015 року [Текст]. – Головне управління статистики у Рівненській області. – Рівне, 2015. – 109 с.

8. Айвазян, С. А. Прикладная статистика. Классификация и снижение размерности [Текст] / С. А. Айвазян, В. М. Бухштабер, И. С. Енюков, Л. Д. Мешалкин. – М.: Финансы и статистика, 1989. – 607 с.

9. Дейвисон, М. Многомерное шкалирование [Текст] / М. Дейвисон. – М.: Финансы и статистика, 1988. – 254 с.

10. Gan, G. Data Clustering: Theory, Algorithms, and Applications [Text] / G. Gan, C. Ma, J. Wu. – ASA-SIAM Series on Statistics and Applied Probability, SIAM, Philadelphia, ASA, Alexandria, VA., 2007. – 489 p. doi: 10.1137/1.9780898718348

## References

1. Targeted Comprehensive Program for Research of NAS of Ukraine “Grid-infrastructure and Grid-technologies for scientific and applied applications” (2013). Resolution of the Presidium of the National Academy of Sciences of Ukraine of 11.12.2013., 164. Available at: <http://www1.nas.gov.ua/infrastructures/Legaltexts/nas/2013/Pages/>

2. Zgurovsky, M. Z. (Ed.) (2010). Analiz staloho rozvytku – hlobal'nyy i rehional'nyy konteksty [Analysis of sustainable development – global and regional contexts]. Global'nyy analiz yakosti ta bezpeky zhyt'tja ljudej. Kyiv: NTUU “KPI”, 252.

3. Kussul, N. M., Shelestov, A. Y. (2008). Grid-sistemy dlja zadach issledovanija Zemli. Arhitektura, modeli i tehnologii [Grid-systems for the tasks of the study of the Earth. Architecture, models and technologies]. Kyiv: Naukova dumka, 452.

4. Zgurovsky, M. Z., Petrenko, A. I., Kiselev, G. D. (2009). Stvorennya natsional'noyi Grid-infrastruktury dlya zabezpechennya naukovykh doslidzhen' [Creation of National Grid-infrastructure for scientific research]. Information Technologies in Education, 4, 12–17.

5. Soshnikova, L. A., Tamashyevych, V. N., Uebe, H., Schafer, M. (1999). Mnomomernyj statysticheskij analiz v ekonomike [Multidimensional statistical analysis in the economy]. Moscow: UNITY – DANA, 598.

6. Backhaus, K., Erichson, B., Plinke, W., Weiber, R. (2006). Multivariate Analysemethoden: Eine anwendungsorientierte Einführung. Springer Berlin, 830.

7. Ekonomichne i socialne stanovysche Rivnenskoji oblasti za sichen'-lypen' 2015 roku : statystychnyj bjuletjen' [The economic and social situation of the Rivne region in January-July 2015 : statistical bulletin] (2015). Department of Statistics in Rivne region. Rivne, 109.

8. Ayvazyan, S. A., Buchstaber, V. M., Enyukov, I. S., Meshalkyn, L. D. (1989). Prikladnaya statistika. Klassifikaciya i snizhenie razmernosti [Applied Statistics. Classification and dimension reduction]. Moscow: Finance and Statistics, 607.

9. Dejvyson, M. (1988). Mnomomernoe shkalyrovanye [Multidimensional Scaling]. Moscow: Finansy i statistika, 254.

10. Gan, G., Ma, C., Wu, J. (2007). Data Clustering: Theory, Algorithms, and Applications. ASA-SIAM Series on Statistics and Applied Probability, SIAM, Philadelphia, ASA, Alexandria, VA., 489. doi: 10.1137/1.9780898718348

*Рекомендовано до публікації д-р техн. наук, професор Бомба А. Я.  
Дата надходження рукопису 22.09.2015*

**Зубик Людмила Володимирівна**, старший викладач, кафедра комп'ютерних наук, Національний університет водного господарства і природокористування, вул. Соборна, 11, м. Рівне, Україна, 33028  
E-mail : labrob@ukr.net

**Гладка Олена Миколаївна**, кандидат технічних наук, старший викладач, кафедра комп'ютерних наук, Національний університет водного господарства і природокористування, вул. Соборна, 11, м. Рівне, Україна, 33028  
E-mail: viklom@ukr.net

**Карпович Иван Миколайович**, кандидат фізико-математичних наук, доцент, кафедра комп'ютерних наук, Національний університет водного господарства і природокористування, вул. Соборна, 11, м. Рівне, Україна, 33028  
E-mail: ivankarp2@ukr.net

**Савич Віра Олексіївна**, кандидат фізико-математичних наук, доцент, кафедра комп'ютерних наук, Національний університет водного господарства і природокористування, вул. Соборна, 11, м. Рівне, Україна, 33028,  
E-mail: sav\_vir2007@ukr.net