

УДК 519.816

**В. В. Циганок, П. Т. Качанов, С. В. Каденко,
О. В. Андрійчук, Г. А. Гоменюк**

Інститут проблем реєстрації інформації НАН України
вул. М. Шпака, 2, 03113 Київ, Україна

Експериментальний аналіз технології експертного оцінювання

Викладено опис і результати експериментального дослідження експертної технології, що створена на базі методу парних порівнянь, в основу якого для більш ефективного отримання інформації покладено надання експертові можливості проводити оцінювання в шкалі, до-кладність якої адекватно відображає його компетентність у питанні експертизи. Метою дослідження є порівняльний аналіз запропонованої технології та наявних, що використовують для експертного оцінювання деяку єдину визначену вербальну шкалу. Результати дослідження, які отримано на основі суджень експертів, залучених до експертизи, свідчать на користь новоствореної технології.

Ключові слова: підтримка прийняття рішень, слабко структурована предметна область, експертне оцінювання, компетентність експертів, шкала експертного оцінювання, експеримент.

Вступ

Для підвищення достовірності експертної інформації, що застосовується для прийняття рішень в слабко структурованих предметних областях передбачається застосування переважно групових експертіз, і тому необхідно є процедура обробки та узагальнення інформації, що отримана від різних експертів. У [1] викладено обґрунтування необхідності застосування для проведення експертіз у таких предметних областях (вони характеризуються так званими невідчутними, англійською мовою — intangible, факторами) методу парних порівнянь [2], як методу відносних вимірювань. Як правило, у процесі парних порівнянь експертам пропонуються на вибір значення з вербальної шкали, адже людина краще оперує вербальними поняттями, аніж чисельними значеннями.

Для оцінки варіантів рішень на основі експертної інформації можуть застосовуватися групові методи парних порівнянь із використанням як ординальних, так і кардинальних експертних оцінок. До таких методів у класі ординального оці-

нювання можна віднести правила Борда [3], Кондорсе [4], медіану Кемені [5] та їхні модифікації. У контексті кардинального оцінювання слід згадати методи, які запропоновано та описано в працях [6–10]. Характерною особливістю вказаних методів є використання в процесі експертного оцінювання (ЕО) певної, апріорно заданої шкали. Це може приводити до неадекватного відображення знань експерта про співвідношення між об'єктами предметної області. Причинами такої неадекватності є: з одного боку — тиск на експерта, який вимушений вибирати оцінки в рамках заданої шкали навіть у випадках, коли він невпевнений у власній оцінці співвідношення між заданою парою об'єктів; з іншого боку — невідповідність ступеня деталізації шкали більш глибоким знанням експерта про співвідношення між об'єктами. Ще одна причина — брак у експерта апріорного уявлення про реальні чисельні співвідношення, яке відповідає вербалним значенням поділок запропонованої шкали. Тобто, кожний експерт із групи може вкладати різний зміст у такі поняття як, наприклад, «сильна перевага» (у фундаментальній шкалі Saati цьому значенню відповідає перевага у 5 разів).

Запропонована технологія експертного оцінювання

Для усунення вищевказаних недоліків запропоновано застосувати підхід, який дозволяє одночасно використовувати в процесі експертизи парні порівняння, що задані в різних шкалах [11]. Пропонується технологія ЕО, яка дозволяє експертові при кожному окремому парному порівнянні вибирати шкалу з кількістю градацій, що адекватно відображає його компетентність у питанні, яке розглядається.

Відповідно до вищезгаданої технології [11] для визначення чисельного відповідника кожному верbalному виразу шкали експерти окремо для введення кожного парного порівняння мають змогу вибрати одну з наступних шкал:

- ординальну;
- ціло-чисельну;
- збалансовану;
- степеневу;
- Ма–Чженга;
- Донегана–Додд–Макмастера.

Дляожної із кардинальних шкал експерт має змогу обирати кількість поділок. У даній технології ЕО кількість поділок шкали можливо обирати в межах від 3 до 9.

У рамках підходу всі оцінки, що задані у вигляді парних порівнянь у різних шкалах, приводяться до уніфікованих значень, причому кожній шкалі, залежно від її ступеня деталізації, присвоюється відповідна вага. Щоб запобігти втраті інформації, приведення оцінок виконується до більш деталізованої шкали із усіх, що використовувалися. Власне, процедура уніфікації значень парних порівнянь, що задані в різних шкалах, описана у [12].

Отримані результати парних порівнянь (у загальному випадку — неповних) агрегуються за допомогою комбінаторного методу [13], який характеризується вищою ефективністю порівняно з іншими методами агрегації [14]. Сутність методу полягає в повному переборі наявних множин, що складаються з мінімальних наборів інформаційно-вагомих елементів матриці парних порівнянь (МПП) з по-

дальнім усередненням ваг, отриманих за кожною такою множиною. Крім того, метод характеризується універсальністю стосовно можливості обробки як повних, так і неповних МПП, що досить корисно при реалізації даної технології ЕО.

Брак у експерта уявлення про реальні чисельні співвідношення, що відповідають поділкам вербальних шкал, значною мірою компенсується інформаційними підказками (див. рис. 3), які реалізовані в інтерфейсі підсистеми вводу інформації.

Обґрунтування необхідності експерименту

Очевидно, що запропонована технологія потребує обґрунтування. Експертні технології, які застосовуються в слабко структурованих предметних областях, де принципово не існує еталонних значень ваг об'єктів, взагалі виключають аналітичне обґрунтування та строгое доведення адекватності.

Отже, єдиним способом перевірки та обґрунтування правомірності використання запропонованої технології є її експериментальне дослідження на багатьох прецедентах ЕО.

У даному випадку імітаційне моделювання експертних оцінок (на зразок того що застосовується у [15, 16]), є неприйнятним, оскільки в рамках експериментального дослідження передбачається, що експерт сам має визначати, наскільки адекватно агреговані результати експертизи відображають його власні уявлення про співвідношення об'єктів. Інформацію про відповідність отриманих співвідношень уявленням експерта можна отримати лише від самого експерта: змоделювати такий показник відповідності неможливо.

Мета і сутність експерименту

Мета експерименту — порівняння запропонованої технології із наявними технологіями отримання відносних ваг факторів, зокрема, із тими, що використовують широко розповсюджене сімейство методів аналізу ієархій і мереж Т. Сааті [6]. Порівнювати пропонується ступінь адекватності уявлень експерта про реальні співвідношення між об'єктами предметної області (факторами) до співвідношень, визначених за допомогою інструментарію тієї чи іншої технології ЕО. Фактично, порівнюються результати, які отримано на основі технології, з еталонами оцінок, що сформувались у свідомості експерта.

Результати (відносні ваги факторів), які отримано на основі технології, що тестиється, порівнюються із вагами, отриманими на основі парних порівнянь у фундаментальній шкалі з 5-ма та 9-ма поділками. Ваги критеріїв (факторів) для парних порівнянь у фундаментальній шкалі визначаються методом власного вектора Т. Сааті на основі МПП.

Експеримент пропонується проводити згідно з нижченаведеною схемою (див. рис. 1), яка включає 4 основних етапи.

1. *Формулювання цілі (проблеми) та факторів, які на неї впливають.* На цьому етапі експерт обирає проблему, в якій він/вона добре розуміється і формулює від 5-ти до 7-ми відносно незалежних факторів, які, на його думку, є найбільш суттєвими для вирішення цієї проблеми.

2. Парні експертні порівняння факторів у трьох різних шкалах. Усі парні порівняння, що мають бути виконані протягом усього експерименту (для всіх трьох технологій) подаються експертами у випадковому порядку.

3. Обчислення векторів ваг факторів. Для визначення векторів відносних ваг об'єктів відповідними методами обробляються МПП, які отримано на попередньому етапі.

4. Ранжирування розрахованих векторів ваг. На цьому етапі експертом проводиться ранжирування трьох розрахованих на третьому етапі експерименту векторів ваг факторів у порядку зменшення відповідності до власного уявлення. Вектори у вигляді стовпчикових діаграм виводяться на екран у випадковому порядку і без зазначення назви.

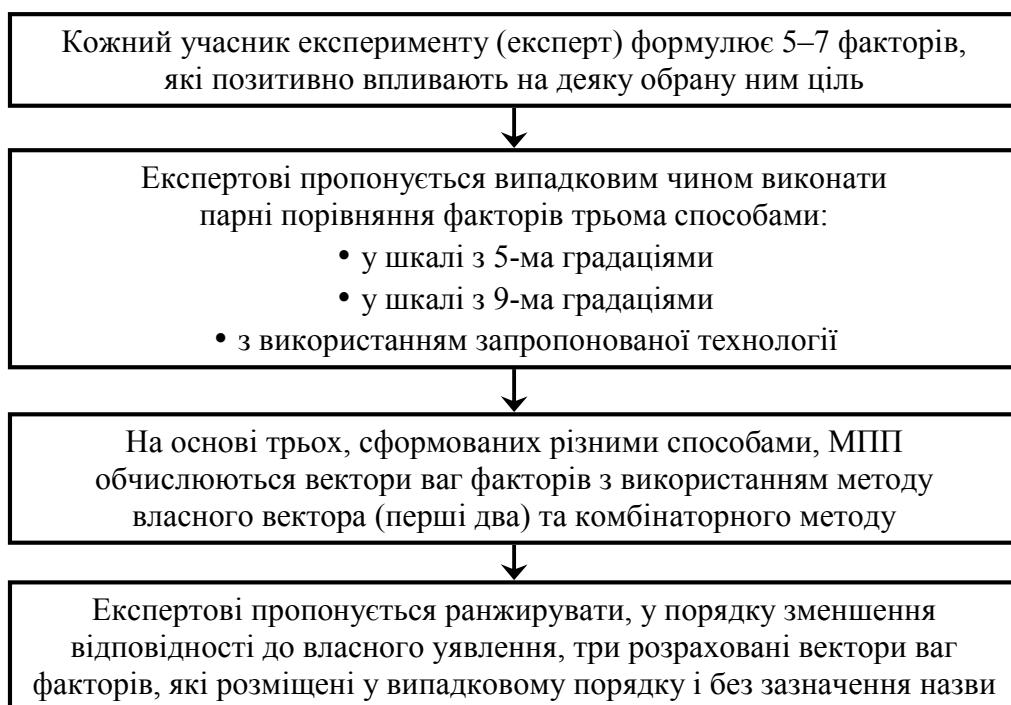


Рис. 1. Основні етапи експериментального дослідження

Кількісним показником, що формується в результаті проведення експерименту є ступінь переваги тієї чи іншої технології ЕО (частота вибору її експертом найкращою у ранжируванні серед тих, що тестувались).

Особливості та засоби забезпечення чистоти експерименту

На першому із узказаних етапів пропонується, щоб експерт сам вибирав предметну область, в якій він вважає себе достатньо компетентним. Експертам пропонується сформулювати проблему, яка, на його погляд, є найбільш зрозумілою для нього. Надання експертам можливості самостійного вибору предметної області для оцінювання гарантує неупередженість організаторів експертизи (адже

предметну область пропонує не організатор експертизи, а сам експерт). Організатор експертизи не нав'язує експертові предметну область, в якій останній може виявитися недостатньо компетентним.

Після формулювання проблеми експертові пропонується сформулювати набір факторів, які її описують. Цей, один і той самий, набір факторів використовується для оцінювання різних технологій ЕО в процесі експерименту — таким чином забезпечується можливість коректного порівняння результатів, що отримані за різними технологіями.

При формулюванні критеріїв (факторів) вимагається, щоб вони найбільш повно описували проблему і при цьому були незалежними між собою (по можливості, «не перекривалися» між собою, не мали значного взаємовпливу). Це є необхідною умовою отримання достовірних результатів за допомогою методів обробки МПП.

Передбачається також можливість довільно задавати порядок подання факторів для ЕО: інтуїтивно експерт згадує (і, відповідно, задає) в першу чергу більш важливі, на його думку, фактори — і цей порядок факторів залишається одним і тим самим для всіх методів. Це є також важливим для забезпечення однакової достовірності результатів, що отримані на основі обробки парних порівнянь (бо в методі парних порівнянь кількість порівнянь, в яких приймає участь конкретний об'єкт, залежить від порядку об'єктів).

Кількість факторів не має перевищувати 7 ± 2 . Ця умова визначається психофізичними обмеженнями середньостатистичної людини (в т.ч., експерта) [17, Міллер]. До того ж, саме вона відіграє вирішальну роль при формуванні діапазонів значень і кількості поділок шкал, що використовуються.

Випадковий вибір пари факторів (об'єктів), які пропонуються експертові для порівняння, а також технології їхнього ЕО, *на другому етапі*, дозволяє зменшити кореляцію між оцінками співвідношень, що отримані різними методами для одних і тих самих об'єктів під час одного і того ж сеансу експерименту. Таким чином, ми отримуємо можливість домогтися взаємної незалежності значень окремих парних порівнянь.

Ранжирування технологій ЕО проводиться «наосліп», тобто, *на четвертому етапі* експертові пропонується за стовпчиковими діаграмами обрати один з трьох векторів ваг факторів. Під жодною з діаграм не вказується тип шкали, у якій задавалися оцінки, що склали основу для побудови вектора ваг. Таким чином, гарантується неупередженість ранжирування технологій ЕО за ступенем відповідності ваг факторів уявленням експерта про проблему.

Обробка та аналіз результатів експерименту

Результат кожного експерименту являє собою файл з наступною інформацією:

- 1) прізвище, ім'я, по батькові експерта;
- 2) назва проблеми;
- 3) перелік факторів;
- 4) МПП факторів, що отримані із застосуванням різних технологій ЕО;
- 5) час, який витрачено експертом на кожне запитання;

- 6) ранжирування технологій ЕО (векторів ваг критеріїв);
- 7) обґрунтування експертом свого ранжирування.

У рамках даного експерименту кожен експерт може брати участь кілька разів поспіль за умови формулювання різних проблем.

Одразу після завершенняожної окремої експертизи її результати перевіряються на адекватність, адже подальшу статистичну обробку слід проводити на основі вибірки адекватних ранжирувань технологій ЕО. Перевірка (очищення, відсів) включає два етапи.

1. При аналізі не враховуються (видаляються) результати, які можна розцінювати як «недбалі». Йдеться про: а) результати, на отримання яких експертом витрачений занадто малий час (адже експертові потрібно принаймні кілька секунд на те, щоб сформулювати обдуману відповідь чи оцінку); б) МПП, у яких співпадає більшість оцінок; в) файли з результатами, що корегувалися вручну; г) неповністю заповнені МПП, у яких за наявними елементами не можна відновити решту елементів (для випадку неповних парних порівнянь [12]).

2. Відсіюються вектори ваг факторів, якщо між відповідними елементами МПП, на основі яких вони отримані, виявлено значні протиріччя (суттєва неузгодженість). Наприклад, коли один з трьох отриманих векторів ваг факторів суттєво відрізняється від інших двох, експерт відкине цей вектор, а відтак, і відповідну технологію ЕО, і не може включити його у кінцеве ранжирування. Навіть, якщо даний вектор (і відповідна технологія) буде включений експертом у ранжирування як найменш адекватний, усе це ранжирування не нестиме інформації про відповідність отриманих векторів ваг факторів уявленням експерта про проблему.

Поетапна процедура виконання експерименту експертом

Для проведення експерименту був програмно реалізований макет робочого місця експерта, що призначено для проведення експертиз з оцінки впливів факторів, які характеризують конкретну проблему, на основі парних порівнянь у вищевказаних шкалах.

Поетапна процедура роботи експерта проілюстрована на рис. 2–5.

На рис. 2 зображено екранну форму для уведення початкових даних для експерименту (див. 1-й етап). На рис. 3 показано приклади інтерфейсу виконання парних порівнянь за допомогою різних технологій ЕО. Слід звернути увагу на наявність інтерактивної підказки, зображені на екранній формі знизу рисунка, яка використовується в технології, що тестиється. Ця підказка наглядно демонструє співвідношення між об'єктами, яке визначає експерт під час порівняння.

У випадку, якщо експерта не влаштовує співвідношення між об'єктами, які відповідають вербальному значенню деякої шкали, йому в рамках запропонованої технології пропонується можливість підібрати найбільш прийнятну для нього шкалу, керуючись підказками, як показано на рис. 4.

На рис. 5 зображено екранну форму інтерфейсу для введення експертом результатуючого ранжирування технологій ЕО з обов'язковим змістовним поясненням свого вибору.

Сформулюйте ціль і 5-7 факторів

Ціль	
Проблема	
5-7 факторів, що позитивно впливають на Ціль	
Фактор 1	
Фактор 2	
Фактор 3	
Фактор 4	
Фактор 5	

Ваше П.І.Б. Іваненко Іван Іванович

Підтвердити

Рис. 2. Екранна форма для формування цілі (проблеми) та низки факторів

Зробіть вибір

Фактор 2	ВПЛИВАЄ	Ніж Фактор 3
<input type="radio"/> Рівнозначні <input type="radio"/> Менше <input checked="" type="radio"/> Більше <input checked="" type="radio"/> Середнє <input type="radio"/> Сильно <input type="radio"/> Дуже сильно <input type="radio"/> Абсолютно		<input checked="" type="checkbox"/> Підтвердити
Фактор 2	ВПЛИВАЄ	Ніж Фактор 5
<input type="radio"/> Рівнозначні <input checked="" type="radio"/> Менше <input type="radio"/> Більше <input type="radio"/> Слабко або незначно <input type="radio"/> Середнє <input type="radio"/> Більше ніж середнє <input type="radio"/> Сильно <input checked="" type="radio"/> Більше ніж сильно <input type="radio"/> Дуже сильно <input type="radio"/> Дуже, дуже сильно <input type="radio"/> Абсолютно		<input checked="" type="checkbox"/> Підтвердити
Фактор 3	ВПЛИВАЄ Більше	Ніж Фактор 4
<input type="radio"/> Менше <input type="radio"/> Слабко або незначно <input type="radio"/> Середнє <input type="radio"/> Сильно <input type="radio"/> Дуже сильно <input checked="" type="radio"/> Абсолютно		

Виберіть типу шкали

Рис. 3. Фрагмент інтерфейсу користувача для виконання парних порівнянь

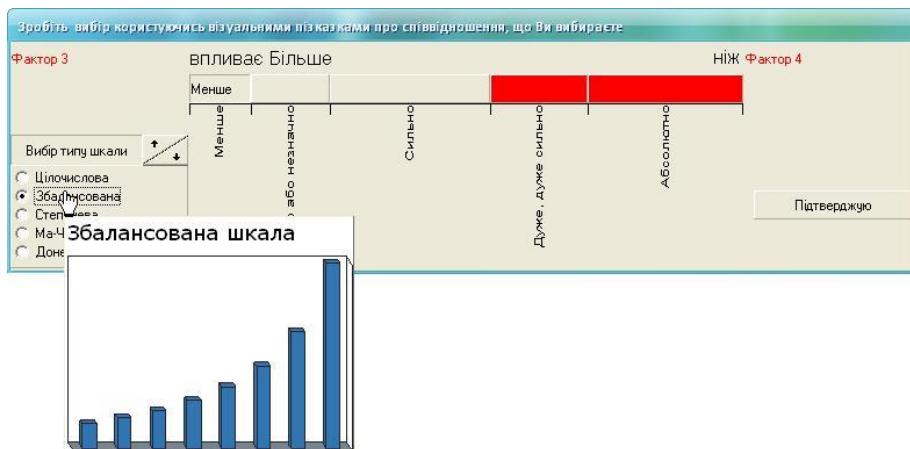


Рис. 4. Інтерфейс вибору шкали

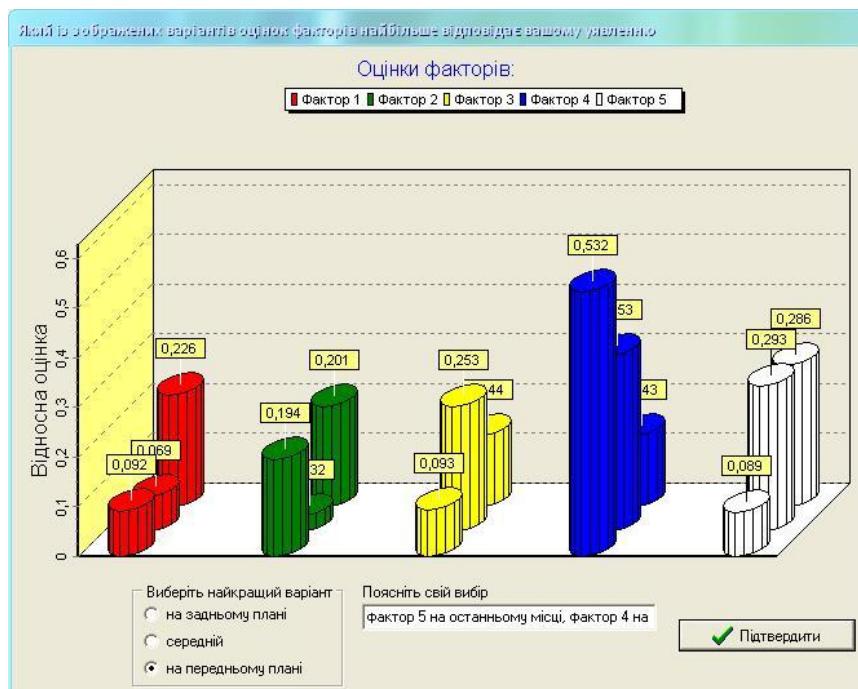


Рис. 5. Екранна форма для введення результуючого ранжирування

Чисельні результати експерименту

Як експерти в експерименті виступали співробітники Лабораторії Систем підтримки прийняття рішень Інституту проблем реєстрації інформації НАН України та студенти-магістранти Державного університету інформаційно-комунікаційних технологій МОНМС України. На даний момент в експерименті взяли участь близько 100 респондентів. Після відсіву вибірка склала 63 ранжирування векторів ваг факторів (і відповідних ним технологій EO), 2 ранжирування виявилися неповними (складаються не з трьох, а з двох векторів). Результати зведені у таблицю.

Назва технології ЕО при застосуванні парних порівнянь	Результати експерименту		
	Кількість респондентів, які присвоїли даній технології ранг		
	«1»	«2»	«3»
У фундаментальній шкалі переваг з 5-ма поділками	10	15	37
У фундаментальній шкалі переваг з 9-ма поділками	12	33	17
Запропонована технологія, що тестиється	41	15	7

Як видно з таблиці, більшість респондентів надала перевагу запропонованій технології ЕО, що ґрунтується на агрегації результатів парних порівнянь, отриманих у різних шкалах [12].

Висновки

Отримані результати експерименту дають змогу зробити низку висновків:

- 1) запропонована у [11, 12] технологія ЕО дозволяє проводити експертизи з оцінки факторів впливу у слабко структурованих предметних областях;
- 2) технологія ЕО, яка використовує різні шкали, має суттєві переваги порівняно з іншими, широко розповсюдженими, технологіями, з якими проводилося порівняння;
- 3) той факт, що більшість респондентів, які брали участь в експерименті, надали запропонованій технології перевагу, вказує на те, що можливість введення оцінок у різних шкалах дійсно дозволяє підвищити ступінь адекватності результатів експертизи;
- 4) експеримент підтверджує доцільність використання запропонованої технології у наявних і нових системах підтримки прийняття рішень.

1. Saaty T.L. Scales from Measurement Not Measurement from Scales! [Electronic resource]: Proceedings of MCDM 2004. — Whistler, B.C., Canada. — Aug. 6-11 / T.L. Saaty // Access mode: <http://www.bus.sfu.ca/events/mcdm/MCDMProgram/Papers/AP168 CF Saaty Scales.pdf> —
2. David H.A. The Method of Paired Comparisons / H.A. Devid // N.Y.: Oxford Univ. Press. — 1988.
3. Borda J.C. Mémoire Sur Les Elections au Scrutin. Historie de L'académie Royale Des Sciences / J.C. Borda. — Paris, 1781. — 657 p.
4. Marquis de Condorcet. Essai Sur L'application de L'analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix: 1785 [Electronic resource] / Marquis de Condorcet. — Access mode: <http://gallica.bnf.fr/ark:/12148/bpt6k417181>
5. Кемени Дж.Г. Кибернетическое моделирование / Дж. Г. Кемени, Дж. Л. Снелл. — М.: Соврадио, 1972. — 192 с.
6. Saaty T.L. The Analytic Hierarchy/Network Process / T.L. Saaty // RACSAM. — 2008. — 102(2). — P. 251–318.

7. *Подиновский В.В.* Парето-оптимальные решения многокритериальных задач / В.В. Подиновский, В.Д. Ногин. — М.: Наука. — 1982. — 255 с.
8. *Литвак Б.Г.* Экспертная информация. Методы получения и анализа / Б.Г. Литвак. — М.: Радио и связь, 1982. — 185 с.
9. *Гнатієнко Г.М.* Експертні технології прийняття рішень / Г.М. Гнатієнко, В.Є. Снітюк // К.: ТОВ «Маклаут», 2008. — 444 с.
10. *Циганок В.В.* Комбінаторний алгоритм парних порівнянь зі зворотним зв'язком з експертом / В.В. Циганок // Реєстрація, зберігання і оброб. даних. — 2000. — Т. 2, № 2. — С. 92–102.
11. *Циганок В.В.* Вибір шкали оцінювання експертом у процесі виконання ним парних порівнянь в системах підтримки прийняття рішень / В.В. Циганок // Реєстрація, зберігання і оброб. даних. — 2011. — Т. 13, № 3. — С. 92–105.
12. *Циганок В.В.* Агрегація групових експертних оцінок, що отримані у різних шкалах // В.В. Циганок // Реєстрація, зберігання і оброб. даних. — 2011. — Т. 13, № 4. — С. 74–83.
13. *Циганок В.В.* Метод обчислення ваг альтернатив на основі результатів парних порівнянь, проведених групою експертів // Реєстрація, зберігання і оброб. даних. — 2008. — Т. 10, № 2. — С. 121–127.
14. *Циганок В.В.* Визначення ефективності методів агрегації експертних оцінок при використанні парних порівнянь / В.В. Циганок // Реєстрація, зберігання і обробка даних. — 2009. — Т. 11, № 2. — С. 83–89.
15. *Saaty T.L.* Decision Making with Dependence and Feedback: The Analytic Network Process / T.L. Saaty. — Pittsburgh: RWS Publications, 1996 — 370 p.
16. *Цыганок В.В.* Имитационное моделирование экспертивных оценок для тестирования методов обработки информации в системах поддержки принятия решений // В.В. Цыганок, С.В. Каденко, О.В. Андрейчук / Проблемы управления и информатики. — 2011. — № 6. — С. 84–94.
17. *Miller G.A.* The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information / George A. Miller // The Psychological Review. — 1956. — 63(2). — P .81–97.

Надійшла до редакції 22.02.2012