

**Л. Ф. Василевич¹, А. Ю. Михайлюк¹,
В. П. Тарасенко², О. К. Тесленко²**

¹Київський університет імені Бориса Грінченка
вул. Маршала Тимошенка, 13-б, 04212 Київ, Україна

²Національний технічний університет України «КПІ»
Прспект Перемоги, 37, 03057 Київ, Україна

Функціонально-орієнтований підхід до проектування інтелектуальних інформаційно-аналітичних систем

Розглянуто універсальний формалізований підхід до організації процедури проектування спеціалізованих інформаційно-аналітичних систем на базі бібліотек уніфікованих вузькофункціональних програмних модулів.

***Ключові слова:** інформаційно-аналітичні системи, проектування, стандартні функціональні профілі.*

Постановка задачі

Важливим інструментом підвищення ефективності будь-якої професійної діяльності в умовах суспільства, що базується на знаннях, є інтелектуальні інформаційно-аналітичні системи (ІАС) відповідної спеціалізації [1, 2]. Однак, незважаючи на високий суспільний попит і значний інженерно-науковий ресурс, спрямований вже протягом тривалого часу на створення програмних засобів даного класу, їхнє впровадження досі не набуло масового характеру. Аналіз ситуації показує, що одним з головних чинників «пробуксовування» у широкому розповсюдженні ІАС є їхня висока вартість (від десятків до сотень тисяч доларів США), котра, в свою чергу, є наслідком унікальності кожної такої розробки [3]. Дійсно, створення універсальних ІАС абсолютно нереальне через величезне розмаїття потенційних сфер і форм їхнього впровадження, а реалізація вузькоспеціалізованої системи, незважаючи на наявність достатньо розвинутої теоретичної бази (апарат OLTP, OLAP, Data Mining тощо [4–6]) і відповідного програмного інструментарію, щоразу починається практично «з нуля» [7–9], значною мірою спрямовується досвідом та інтуїцією розробників і при цьому наперед не гарантує якісного результату через відсутність ефективної спеціальної методології проектування. Таким чином, створення формалізованих підходів до проектування ІАС, за своєю структурно-алгоритмічною організацією оптимально орієнтованих на конкретне застосування, являє собою актуальну науково-практичну задачу.

Метою даної статті є розробка формалізованого підходу до проектування високо-ефективних функціонально-орієнтованих (тобто, спрямованих на конкретний об'єкт впровадження) інтелектуальних інформаційно-аналітичних систем, спроможного суттєво прискорити і знизити вартість процесу їхньої розробки та реалізації.

Формування проектів інтелектуальних інформаційно-аналітичних систем на базі бібліотеки уніфікованих програмних модулів

Аналіз типових напрямків застосування ПАС, варіантів їхньої експлуатації користувачами, особливостей архітектури та технологічного виконання, а також інших чинників дозволяють сформуванню їхню класифікацію [10], подавши її множиною вигляду $\Lambda = \{\lambda_l : l = \overline{1, L}\}$, де λ_l — деякий l -й клас ПАС; L — кількість «листоків» на дереві класифікації, тобто класів ПАС.

Найбільш вагомою відмінністю між ПАС різних класів є перелік елементарних технологічних задач, для розв'язання котрих призначені системи відповідного класу. В загальному випадку прикладами таких задач, характерних для ПАС освітнього спрямування, є семантичний пошук електронних текстових інформаційних об'єктів, класифікація електронних текстових інформаційних об'єктів за тематикою, машинний переклад тощо. Введемо поняття множини характерних для ПАС технологічних задач

$$\Gamma = \{\gamma_g : g = \overline{1, G}\},$$

де γ_g — деяка g -та елементарна технологічна задача; G — загальна кількість технологічних задач, для розв'язання котрих призначені ПАС. Як витікає з попереднього, існує відображення $f^s : \Lambda \rightarrow \Gamma$, причому воно багатозначне і неін'єктивне, тобто ПАС кожного класу, як правило, призначені для розв'язання кількох елементарних технологічних задач, і кожна така задача може бути характерною для систем різних класів.

Назвемо функціональною послугою ПАС набір функцій, що забезпечує системі можливість розв'язання однієї елементарної технологічної задачі [11]. Тоді можна стверджувати, що множина всіх функціональних послуг ПАС $\Psi = \{\psi_m : m = \overline{1, G}\}$, де ψ_m — деяка m -та функціональна послуга, є областю значень деякого відображення $f^c : \Gamma \rightarrow \Psi$, причому f^c бієктивне за визначенням. Тоді суперпозиція двох відображень $f^d = f^s \circ f^c$, також є багатозначною і неін'єктивною та $f^d : \Lambda \rightarrow \Psi$. Отже, кожному класу ПАС λ_l з Λ відповідає певна множина функціональних послуг, які реалізуються системами класу λ_l , причому інформаційно-аналітичні системи різних класів можуть використовувати одні й ті самі послуги. Позначимо

$$\Psi_l^k = f^d(\lambda_l)$$

і назвемо множиною функціональних послуг ПАС l -го класу, причому

$$\Psi_l^k = \{\psi_{ln} : n = \overline{1, N_l} \wedge \psi_{ln} \in \Psi\},$$

де ψ_{ln} — деяка n -та функціональна послуга, характерна для ПАС l -го класу, а N_l — загальна кількість таких послуг.

У загальному випадку відповідність системи деякому класу λ_l з Λ може бути досягнута за рахунок реалізації різних сукупностей функціональних послуг [11–13]. Одним з чинників цього явища є, зокрема, взаємозамінність деяких послуг у більшості

практичних застосувань, наприклад, взаємозамінність функціональних послуг автоматичного реферування та автоматичного анотування в тих випадках, коли відповідна ПАС повинна забезпечити можливість отримання стислої інформації щодо змісту електронного текстового інформаційного об'єкта. Тобто системи, які відносяться до одного класу, в загальному випадку можуть мати різний набір функцій. Таким чином, якщо функціональність деякої p -ї ($p = \overline{1, P_l}$) системи l -го класу (P_l — загальна кількість систем даного класу) позначити через

$$\Phi_{lp}^u = \{\varphi_{lpq}^u : q = \overline{1, Q_p} \wedge \varphi_{lpq}^u \in \Psi_l^\kappa\},$$

то за визначенням

$$\Phi_{lp}^u \subset \Psi_l^\kappa.$$

Отже, якщо назвати Φ_{lp}^u функціональним профілем деякої p -ї системи l -го класу, а φ_{lpq}^u — послугами, котрі утворюють цей профіль (Q_p — кількість послуг у відповідному профілі), то можна стверджувати, що кожен клас ПАС може мати кілька, в загальному випадку взаємозамінних, функціональних профілів [12, 13] і

$$\bigcup_{p=1}^{P_l} \Phi_{lp}^u = \Psi_l^\kappa.$$

З метою тиражування технічних рішень, уніфікації ПАС тощо функціональний профіль деякої r -ї ($r = \overline{1, R_l}$, причому $R_l \leq P_l$) системи l -го класу може бути прийнятим за офіційний стандарт

$$\Phi_{lr}^c = \{\varphi_{lrs}^c : s = \overline{1, S_r} \wedge \varphi_{lrs}^c \in \Psi_l^\kappa\} \quad (1)$$

для систем відповідного класу. Будемо називати Φ_{lr}^c r -м стандартним функціональним профілем l -го класу ПАС, при цьому φ_{lrs}^c — деяка s -та функціональна послуга у його складі, а S_r — кількість послуг у відповідному профілі. Тоді

$$\Pi_l^c = \{\Phi_{lr}^c : r = \overline{1, R_l}\}$$

є множиною стандартних функціональних профілів l -го класу, де R_l — кількість стандартних функціональних профілів відповідного класу.

У загальному випадку перелік стандартних функціональних профілів для кожного класу систем повинен формуватися апіорі з метою формалізації процедури проектування. При цьому, здебільшого, розробники мають досягти максимально можливої повноти Π_l^c . Враховуючи це, а також (1), можна стверджувати, що

$$\Pi_l^c \subset \Psi_l^\kappa. \quad (2)$$

Множиною всіх стандартних функціональних профілів ПАС назвемо

$$\Pi = \bigcup_{l=1}^L \Pi_l^c.$$

Беручи до уваги (2), можна стверджувати, що

$$\exists f^{\text{жс}} (f^{\text{жс}} : \Lambda \rightarrow \Pi \wedge \Pi_l^c = f^{\text{жс}}(\lambda_l)), \quad (3)$$

причому відображення $f^{\text{жс}}$ є багатозначним. Це означає, що кожному класу ПАС відповідає деяка множина стандартних функціональних профілів. Якщо трактувати функціональний профіль як множину послуг, мінімально достатніх для забезпечення функціональності інформаційно-аналітичних систем певного класу, то це відображення є ін'єктивним, тобто кожен функціональний профіль відповідає тільки одному класу ПАС.

Для потреб програмної реалізації ПАС може бути розроблено бібліотеку програмних модулів $\Theta = \{\theta_t : t = \overline{1, T}\}$, де θ_t — деякий t -й програмний модуль; T — загальний обсяг бібліотеки. Враховуючи наявність у розпорядженні сучасних розробників широкого спектра новітніх технологій із забезпечення інтегрованості та формуванню багатокомпонентного програмного середовища (використання універсальних протоколів обміну, створення шару підпрограмного забезпечення, реалізація агентоорієнтованого підходу тощо [14, 15]) будемо вважати витрати по формуванню системи на базі уніфікованих модулів (інтегрування, налагодження, компонування інтерфейсу користувача тощо) такими, що ними можна знехтувати.

Кожен t -й модуль θ_t має свій функціональний профіль

$$\Phi_t^m = \{\varphi_{th}^m : \varphi_{th}^m \in \Psi \wedge h = \overline{1, H_t}\},$$

де φ_{th}^m — деяка h -та функціональна послуга відповідного профілю, а H_t — кількість функціональних послуг у профілі. Множиною функціональних профілів програмних модулів бібліотеки Θ назвемо

$$\Pi^m = \{\Phi_t^m : t = \overline{1, T}\}.$$

При цьому для ідеальної бібліотеки Θ справедливе твердження, що

$$\bigcup_{t=1}^T \Phi_t^m = \Psi.$$

Це означає, що засобами відповідної бібліотеки може бути реалізовано ПАС будь-якого класу.

Введемо відображення $f : \Theta \rightarrow \Pi^m$, яке, за визначенням, є бієктивним. У загальному випадку з метою уникнення надлишковості програмного коду і для зменшення

ймовірності міжмодульних конфліктів необхідно мінімізувати потужність H_i функціонального профілю пересічного t -го модуля, тобто в ідеалі кожен модуль бібліотеки Θ повинен реалізувати тільки одну функціональну послугу. Виняток доцільно зробити лише для тих випадків, коли деяка послуга ψ_i для своєї реалізації вимагатиме функціональності, котра забезпечується іншою послугою ψ_j , де $1 \leq i \leq G$, $1 \leq j \leq G$, $i \neq j$ (наприклад, реалізація функціональної послуги семантичного пошуку вимагає наявності засобів морфологічного аналізу слів та ряду інших послуг). Послугу ψ_i у цьому випадку будемо називати функціонально похідною від послуги ψ_j , увівши для подібної ситуації позначення відповідно $\psi_i = F(\psi_j)$. Однак, якщо через наявність однієї або кількох функціонально похідних послуг потужність відповідного модуля перевищує одиницю, тоді всі без винятку функціональні послуги, реалізовані у даному модулі, на рівні архітектури доцільно виконувати доступними через його зовнішній інтерфейс з метою максимально повного використання функціональності. Таким чином, можна стверджувати, що

$$\exists \theta_t \exists \theta_u (1 \leq t \leq T \wedge 1 \leq u \leq T \wedge u \neq t \wedge |f(\theta_t) \cap f(\theta_u)| \neq 0), \quad (4)$$

тобто функціональні профілі окремих модулів можуть перетинатися. Причиною тому окрім вищезгаданої функціональної взаємозалежності послуг може слугувати версійність, пов'язана зі звичайною практикою оновлення програмного забезпечення, а також загалом позитивне явище присутності на відповідному сегменті ринку кількох паралельних лінійок програмних продуктів від різних розробників та ряд інших чинників.

Назвемо v -м проектом p' -ї ПАС

$$\Omega_{p'v} = \{\omega_{p'vz} : z = \overline{1, Z} \wedge \omega_{p'vz} \in \Theta \wedge \exists \Phi_{lr}^c \in \Pi_l^c (1 \leq l \leq L \wedge 1 \leq r \leq R_l \wedge \bigcup_{z=1}^Z f(\omega_{p'vz}) \supset \Phi_{lr}^c)\}, \quad (5)$$

де $\omega_{p'vz}$ — деякий z -й програмний модуль у складі p' -ї ПАС згідно v -го проекту; Z ($Z \leq S_r$) — кількість програмних модулів у складі p' -ї ПАС за тим же проектом. Доцільно зауважити, що (5) відображає найбільш типову ситуацію, коли згідно проекту до складу ПАС, що розробляється, включають набір програмних модулів, який мінімально достатній для реалізації системи відповідного класу згідно одного зі стандартних функціональних профілів. Однак на практиці можливі відхилення від стандартних функціональних профілів, як правило, у бік їхнього розширення. Цілком імовірна також розробка спеціальних нестандартних функціональних профілів для окремих систем [13]. Отже, враховуючи останнє зауваження можна стверджувати, що (5) дозволяє сформулювати для довільної p' -ї ПАС, що є об'єктом розробки, деяку множину проектів $A_{p'}$.

При цьому можна ввести відображення $f^n : \Pi_l^c \rightarrow A_{p'}$, для якого $\Omega_{p'v} = f^n(\Phi_{lr}^c)$

Відображення f^n багатозначне і неін'єктивне, тобто, з огляду на (4), окремий функціональний профіль можна реалізувати кількома різними проектами і, внаслідок того, що

$$\exists \theta_t (1 \leq t \leq T \wedge \theta_t \in \Theta \wedge |f(\theta_t)| > 1),$$

цілком імовірний випадок, коли один проект може відповідати кільком функціональ-

ним профілям.

Неважко помітити, що за наявності адекватної класифікації ПАС, широкого переліку стандартних функціональних профілів ПАС, достатньо повної бібліотеки уніфікованих програмних модулів, призначених для розв'язку елементарних технологічних задач ПАС і докладного переліку функціональних профілів уніфікованих програмних модулів вищезгаданої бібліотеки, процедура проектування ПАС на формальному рівні спрощено може бути зведена до реалізації суперпозиції

$$f^p : f^{жс} \circ f^n. \quad (6)$$

З огляду на багатозначність $f^{жс}$ і f^n можна стверджувати, що відображення f^p також є багатозначним. Це означає, що

$$\exists \lambda_l \exists \Phi_{lr}^c (1 \leq l \leq L \wedge 1 \leq r \leq R_l \wedge \Phi_{lr}^c = f^{жс}(\lambda_l) \wedge |f^n(\Phi_{lr}^c)| > 1),$$

тобто згідно (6) в результаті виконання проектування може бути отримана деяка, ймовірно більша за одиницю, кількість проектів, кожен з яких відповідає вимогам щодо наявності необхідних функціональних можливостей. Для завершення процедури проектування ПАС необхідно провести аналіз альтернативних проектів ПАС та здійснити остаточний вибір одного з них. Доречно зазначити, що крім функціонального профілю кожний програмний модуль характеризується також цілим рядом інших параметрів як технічного характеру (наприклад, функціональність, продуктивність, надійність, сумісність тощо), так і «нетехнічного» (вартість, рівень підтримки, ергономічність інтерфейсу користувача, характер ліцензії щодо використання, репутація розробника, відгуки попередніх користувачів тощо) [7–9]. Очевидно, що вибір проекту має бути здійснено з врахуванням наведених характеристик.

Аналіз альтернатив та остаточний вибір проекту інтелектуальної інформаційно-аналітичної системи

Задача аналізу та вибору проектів ПАС має наступні особливості, які обумовлюють її складність:

- багатокритеріальність задачі;
- принципова неможливість отримання ні аналітичними, ні статистичними методами функціональної залежності загального показника ефективності системи (надалі будемо називати загальним показником x системи) від часткових показників ефективності системи (надалі будемо називати частковими показниками w_k системи, причому $k = \overline{1, K}$, де K — кількість часткових показників);
- неможливість кількісно вимірювати окремі часткові показники, або складність їхнього отримання, що обумовлює неповноту та неточність вихідних даних;
- необхідність врахування як інтересів особи, котра приймає рішення (ОПР) про вибір проекту ПАС, так і рівня інформованості цієї особи;
- необхідність оперативно приймати рішення щодо вибору проекту ПАС.

Перелічені особливості роблять практично неможливим кількісний аналіз альтернативних проектів ПАС за допомогою методів теорії ймовірностей та математичної статистики. Більш перспективним для розв'язання цієї задачі вважається апарат теорії

нечітких множин [16, 17]. Далі пропонується методика аналізу та вибору проекту ПАС на основі лінгвістичних змінних [17].

Нехай маємо множину $A_{p'} = \{\Omega_{p',v}; v = \overline{1, N^p}\}$ можливих проектів p' -ї ПАС, кожен елемент якої характеризується загальним показником x у вигляді лінгвістичної змінної (ЛЗ), котра визначається кортежем $\langle E, E_j, j = \overline{1,5}; x \in [0;1] ; \mu_{E_j}(x) \in [0;1] \rangle$, де E — назва ЛЗ (в даній задачі E — це «Загальний показник»); $E_j, j = \overline{1,5}$ — терми ЛЗ, областю значень якої є множина x ; $\mu_{E_j}(x_v)$ — функція належності значення загального показника v -го проекту до терму E_j .

Термами ЛЗ «Загальний показник» можуть бути: E_1 — дуже низьке значення (ДНЗ); E_2 — низьке значення (НЗ); E_3 — середнє значення (СЗ); E_4 — високе значення (ВЗ); E_5 — дуже високе значення (ДВЗ).

Оскільки кількісної характеристики загального показника не існує, можна задавати область його значень інтервалом $[0; 1]$. Використовуючи шкалу Харінгтона і трапецієвидні функції належності термів [18], ЛЗ «Загальний показник» можна задавати наступним чином:

$$E_1 = \langle 0:0:0,1:0,2 \rangle; E_2 = \langle 0,1:0,2:0,3:0,4 \rangle; E_3 = \langle 0,3:0,4:0,5:0,6 \rangle; \\ E_4 = \langle 0,5:0,6:0,7:0,8 \rangle; E_5 = \langle 0,7:0,8:1:1 \rangle.$$

Трапецієвидну функцію належності повністю визначають чотири числа $\langle a:b:c:d \rangle$ [18]. У цій же роботі показано, що подання бокових гілок функцій належності відрізками прямих ліній не знижує загальності задачі оцінки, але суттєво спрощує математичні операції над нечіткими величинами. При цьому ліва та права гілки лінійної функції належності мають вигляд відповідно:

$$\mu(x) = \frac{x-a}{b-a}; x \in [a;b], \quad (7)$$

$$\mu(x) = \frac{d-x}{d-c}; x \in [c;d]. \quad (8)$$

Аналогічно визначають часткові показники кожного проекту w_j . Кількість термів часткових показників повинна дорівнювати кількості термів проекту ПАС. При цьому можлива ситуація, коли різні показники мають різну вагу для ОПР, що можна врахувати за допомогою коефіцієнтів пріоритетності $B_k, k = \overline{1,K}$. Для знаходження усіх коефіцієнтів застосовують методи Фішберна, Уея, Сааті та інші [18].

При розрахунку загального показника проекту спочатку знаходять часткові показники $w_k, k = \overline{1,K}$. Потім для кожного w_k визначають за формулами (7) та (8) значення функцій належності за кожним термом відповідних ЛЗ $\mu_{k,j}(w_k); j = \overline{1,5}$. Далі для кожного однакового за змістом терму (ДНЗ, НЗ, СЗ, ВЗ, ДВЗ) отримуємо зважену оцінку

$$x_j = \sum_k B_k \mu_{kj}(w_k). \quad (9)$$

Загальний показник проекту ПАС знаходять за формулою:

$$x = \sum_j x_j E_{\alpha_j} = \sum_j \sum_k B_k \mu_{kj}(w_k) E_{\alpha_j}, \quad (10)$$

де $E_{\alpha_j} = \frac{E_{\alpha_{j+}} + E_{\alpha_{j-}}}{2}$ — середина α -зрізу нечіткого терму E_j ЛЗ «Загальний показник» [18]; $E_{\alpha_{j+}}$ і $E_{\alpha_{j-}}$ — ліва і права границі α -зрізу терму E_j . При симетричних бокових гілках функції належності E_{α_j} середина α -зрізу терму E_j дорівнює середині ядра E_j ($\alpha = 1$). Для вищенаведених даних маємо:

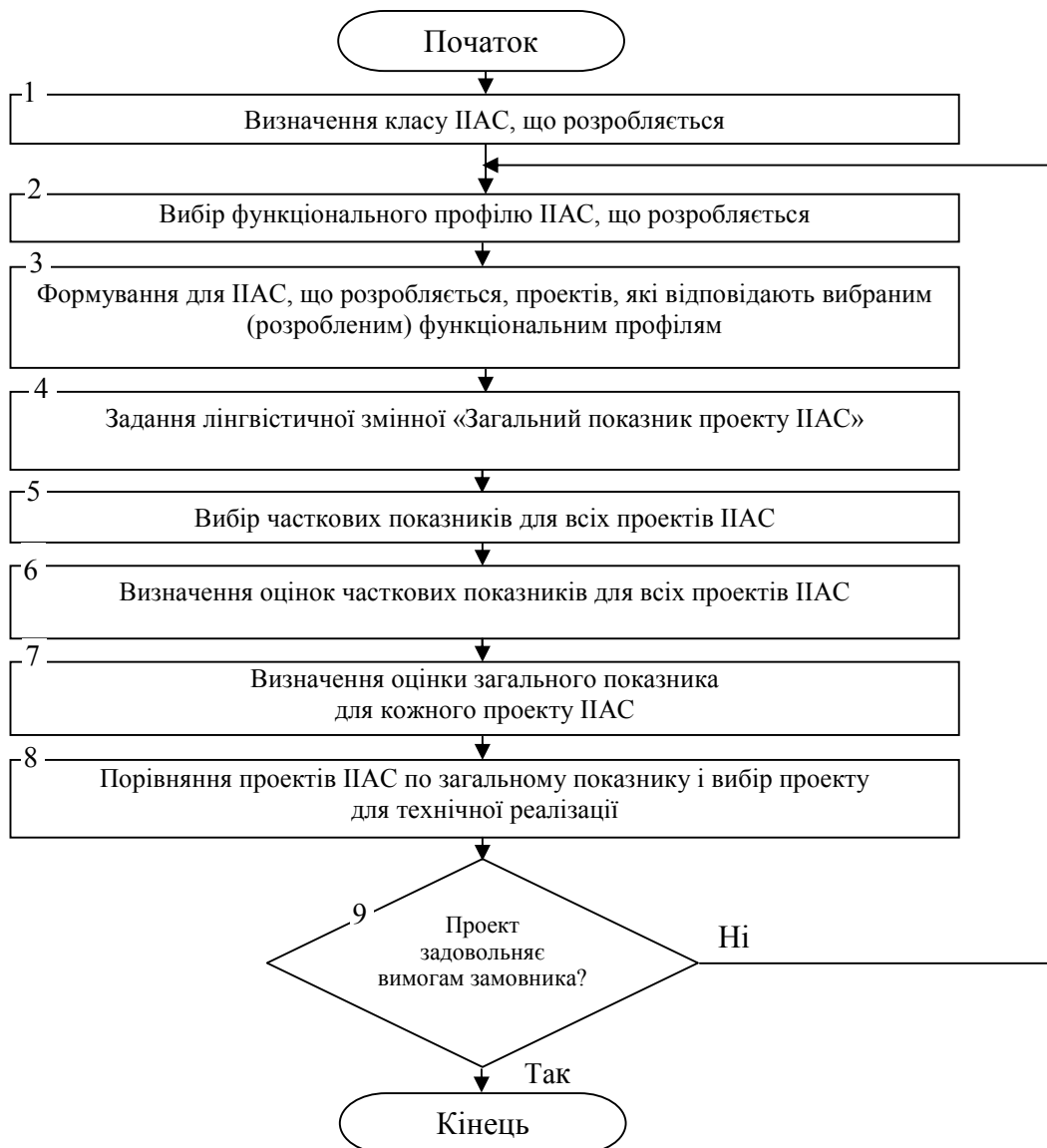
$$E_{\alpha_1} = \frac{0+0,1}{2} = 0,05; \quad E_{\alpha_2} = \frac{0,2+0,3}{2} = 0,25; \quad E_{\alpha_3} = \frac{0,4+0,5}{2} = 0,45;$$

$$E_{\alpha_4} = \frac{0,6+0,6}{2} = 0,65; \quad E_{\alpha_5} = \frac{0,8+1}{2} = 0,9.$$

За загальним показником проводять порівняння різних проектів ПАС. Крім того, загальний показник дозволяє встановити, яким термом характеризується відповідний проект. Коли оцінка за якимось показником є нечіткою величиною, слід проводити дефазифікацію відповідних оцінок. При застосуванні методу дефазифікації по центру ваги маємо:

$$x_{цв} = \frac{\int_a^d x \mu(x) dx}{\int_a^d \mu(x) dx} = (d^2 + c^2 + dc - a^2 - b^2 - ab) / 3(d - a + c - b). \quad (11)$$

Алгоритм проектування ПАС з вибором варіантів на основі ЛЗ приведений на рисунку. Тут один чи декілька стандартних функціональних профілів вибирають (блок 2 алгоритму) з числа тих, які відповідають класу, визначеному для ПАС, що проектується. Можлива розробка спеціального функціонального профілю або модифікація стандартних. У ході задання ЛЗ «Загальний показник проекту ПАС» (блок 4) утворюють всі терми і відповідні функції належності. При виборі часткових показників для всіх проектів ПАС (блок 5) задають відповідні ЛЗ та обчислюють коефіцієнти пріоритетності. Оцінки часткових показників проектів ПАС (блок 6) визначають шляхом експертного оцінювання, розрахунків та аналізу виходячи зі значень показників уніфікованих програмних модулів, які входять до складу даного проекту.



Алгоритм процедури проектування інтелектуальної інформаційно-аналітичної системи

Приклад

Для ілюстрації запропонованого підходу до проектування ПАС розглянемо наступний приклад. Нехай ведеться розробка ПАС, плани щодо застосування якої та інші міркування дозволяють віднести її до класу «12. Контекстно-пошукові ПАС». Припустимо, для цього класу систем передбачено два стандартні функціональні профілі:

$$\Phi_{12.1}^c = \{\psi_7, \psi_{19}, \psi_{27}\}, \quad \Phi_{12.2}^c = \{\psi_7, \psi_{19}, \psi_{31}\},$$

де ψ_7 — функціональна послуга «Контекстний пошук»; ψ_{19} — функціональна послуга «Ранжування пошукового відгуку за релевантністю»; ψ_{27} — функціональна послуга «Автоматичне реферування тексту»; ψ_{31} — функціональна послуга «Автоматичне анування тексту».

Нехай розробникам системи доступна бібліотека уніфікованих програмних модулів, котра серед інших містить, зокрема, наступні модулі: θ_4 з функціональним профілем $\Phi_4^m = \{\psi_7\}$; θ_9 з функціональним профілем $\Phi_9^m = \{\psi_7, \psi_{16}\}$; θ_{41} з функціональним профілем $\Phi_{41}^m = \{\psi_{19}\}$; θ_{45} з функціональним профілем $\Phi_{45}^m = \{\psi_{27}\}$; θ_{68} з функціональним профілем $\Phi_{68}^m = \{\psi_{31}\}$.

Тут ψ_{16} — функціональна послуга «Автоматична орфокоорекція пошукового запиту». Тоді згідно запропонованої методики для реалізації системи, що розробляється, може бути сформовано 4 наступних проекти:

$$\Omega_1 = \{\theta_4, \theta_{41}, \theta_{45}\}, \Omega_2 = \{\theta_9, \theta_{41}, \theta_{45}\}, \Omega_3 = \{\theta_4, \theta_{41}, \theta_{68}\}, \Omega_4 = \{\theta_9, \theta_{41}, \theta_{68}\}$$

(оскільки даний приклад ілюструє процедуру проектування лише однієї ПАС, індекс p' в умовному позначенні проекту є зайвим, тому тут і далі обмежуватимемося лише індексом v).

Для вибору оптимального проекту слід врахувати часткові показники модулів, наведені в табл. 1 (в табл. 1 і далі букву Д після скорочень ДВЗ, ВЗ і т.д. слід читати як «з достовірністю»). Значення деяких часткових показників подано в умовних одиницях, решту, перш за все інтегральні показники, охарактеризовано нечіткими експертними оцінками [18–21] для ЛЗ з п'ятьма термами і діапазоном зміни від 0 до 1. Границі відповідних функцій належності задані за аналогією із загальним критерієм.

Таблиця 1. Параметри уніфікованих програмних модулів

Модуль	Параметр	Значення параметра
θ_4	Швидкість пошуку	ДВЗ Д 0.8 .
	Точність пошуку	ДВЗ Д 0.8 .
	Повнота пошуку	ВЗ Д 1 .
	Надійність	ВЗ Д 0.7 і ДВЗ Д 0.3 .
	Вартість	1800 умовних одиниць
	Ліцензійні умови	СЗ Д 1
θ_9	Швидкість пошуку	ВЗ Д 0.9 .
	Точність пошуку	ВЗ Д 0.9 і ДВЗ Д 0.1 .
	Повнота пошуку	ДВЗ Д 0.6 .
	Надійність	СЗ Д 0.6 і ВЗ Д 0.4 .
	Вартість	1400 умовних одиниць.
	Ліцензійні умови	НЗ Д 1.0 .
θ_{41}	Якість ранжування	ВЗ Д 0.4 і ДВЗ Д 0.6 .
	Надійність	ВЗ Д 0.6 і ДВЗ Д 0.4 .
	Вартість	1000 умовних одиниць.
	Ліцензійні умови	НЗ Д 0.4 і СЗ Д 0.6 .
θ_{45}	Якість реферування, анотування	ВЗ Д 0.7 і ДВЗ Д 0.3 .
	Надійність	СЗ Д 0.9 і НЗ Д 0.1 .
	Вартість	2200 умовних одиниць.
	Ліцензійні умови	ДВЗ Д 0.6 .
θ_{68}	Якість реферування, анотування	ДВЗ Д 0.8 .
	Надійність	СЗ Д 0.9 і НЗ Д 0.1 .
	Вартість	1800 умовних одиниць.
	Ліцензійні умови	НЗ Д 1

Серед проектів $\Omega_1, \Omega_2, \Omega_3, \Omega_4$ слід вибрати оптимальний, враховуючи, що найважливішим елементом системи є модуль контекстного пошуку, тобто θ_4 і θ_9 для відповідних проектів. Порівняння проектів проводитимемо, приймаючи, що узагальнений показник якості кожного з них визначатиметься на основі наступних часткових показників (поданих у порядку зниження пріоритету): швидкість пошуку w_1 ; точність пошуку w_2 ; повнота пошуку w_3 ; якість реферування/анотування w_4 ; якість ранжування w_5 ; гарантоздатність w_6 (інтегральна характеристика, яка визначається надійністю модулів та якістю їхнього ліцензійного супроводу розробником); вартість проекту w_7 .

При цьому вартість проекту будемо вважати рівною сумі вартостей модулів, з яких він складається. Для тих проектів, у яких тільки один модуль визначає якийсь показник, значення цього показника для всього проекту співпадає зі значенням показника для модуля. Для інших показників їхньої оцінки для всього проекту також відповідають значенням ЛЗ в діапазоні від 0 до 1, і для їхнього знаходження потрібно застосувати ту ж саму методику, що і для знаходження загального показника проекту. У цьому разі за допомогою коефіцієнтів пріоритетності можна враховувати, що різні модулі вносять різний внесок у цей показник.

Таким чином, функції належності термів ЛЗ 1–6 показників задамо наступними ($i = \overline{1,6}$): $w_{i1} = \langle 0; 0; 0,1; 0,2 \rangle$, що відповідає ДНЗ; $w_{i2} = \langle 0,1; 0,2; 0,3; 0,4 \rangle$, що відповідає НЗ; $w_{i3} = \langle 0,3; 0,4; 0,5; 0,6 \rangle$, що відповідає СЗ; $w_{i4} = \langle 0,5; 0,6; 0,7; 0,8 \rangle$, що відповідає ВЗ; $w_{i5} = \langle 0,7; 0,8; 1; 1 \rangle$, що відповідає ДВЗ.

Функції належності термів ЛЗ показника 7 дорівнюють: $w_{71} = \langle 0; 0; 2; 3 \rangle$, що відповідає ДВЗ; $w_{72} = \langle 2; 3; 3,5; 4 \rangle$, що відповідає ВЗ; $w_{73} = \langle 4; 4,2; 4,5; 5 \rangle$, що відповідає СЗ; $w_{74} = \langle 4,5; 5; 5,3; 5,5 \rangle$, що відповідає НЗ; $w_{75} = \langle 5,3; 5,5; \infty; \infty \rangle$, що відповідає ДНЗ.

Знайдемо значення показника гарантоздатності w_6 . З цієї метою спочатку для кожного проекту потрібно визначити їхню надійність (позначимо як w_8) та якість ліцензійного супроводу розробниками модулів (позначимо як w_9).

Для кожного однакового за змістом (ДНЗ, НЗ, СЗ, ВЗ, ДВЗ) терму одержимо оцінки:

$$x_j = \sum_k B_k \mu_{kj}(w_k).$$

Для прикладу коефіцієнти B_i отримаємо за методом Фішберна:

$$B_k = \frac{2(T+1-l_k)}{T*(T+1)}, \quad (12)$$

де l_k — номер k -го модуля в порядку його важливості за відповідним показником для проекту; T — загальна кількість модулів в проекті.

Нехай модуль θ_4 більш важливий для показника w_8 ніж модуль θ_{45} , а модуль θ_{45} — ніж модуль θ_{41} . За формулою (12) обчислимо відповідні коефіцієнти пріоритетності для модулів:

$$B_1 = \frac{2(3+1-1)}{3*4} = \frac{1}{2}; \quad B_2 = \frac{1}{6}; \quad B_3 = \frac{1}{3}.$$

Тоді для проекту Ω_1 маємо наступні дані (табл. 2).

Таблиця 2. Дані для розрахунку показника надійності першого проекту

Модуль	Терми показника w_8				
	ДНЗ	НЗ	СЗ	ВЗ	ДВЗ
θ_4	$\mu_{11} = 0$	$\mu_{12} = 0$	$\mu_{13} = 0$	$\mu_{14} = 0,7$	$\mu_{15} = 0,3$
θ_{41}	$\mu_{21} = 0$	$\mu_{22} = 0$	$\mu_{23} = 0$	$\mu_{24} = 0,6$	$\mu_{25} = 0,4$
θ_{45}	$\mu_{31} = 0$	$\mu_{32} = 0,1$	$\mu_{33} = 0,9$	$\mu_{34} = 0$	$\mu_{35} = 0$
$x_j = \sum_{i=1}^3 B_i \mu_{ij}$	0	0,0333	0,3	0,45	0,216

Оцінку узагальненого показника надійності w_8 для проекту Ω_1 знаходимо за формулою (10):

$$w_{18} = \sum_{j=1}^5 x_j E_{\alpha_j} = 0 * 0,05 + 0,0333 * 0,25 + 0,3 * 0,45 + 0,45 * 0,65 + 0,216 * 0,9 = 0,633.$$

При $w_{18} = 0,633$ надійність проекту Ω_1 відповідає терму ВЗ з функцією належності $\mu_4(0,633) = 1$, а всім іншим термам — з функцією належності, рівною нулю.

У такий же спосіб знаходимо показники надійності для чотирьох інших проектів. Для проекту Ω_2 маємо:

$$w_{28} = \sum_{j=1}^5 x_j E_{\alpha_j} = 0 * 0,05 + 0,0333 * 0,25 + 0,6 * 0,45 + 0,3 * 0,65 + 0,067 * 0,9 = 0,533.$$

Для проекту Ω_3 — $w_{38} = 0,672$, а для проекту Ω_4 — $w_{48} = 0,586$.

Аналогічно знаходимо значення показників якості ліцензійного супроводу модулів розробниками, а саме: $w_{19} = 0,554$; $w_{29} = 0,451$; $w_{39} = 0,39$; $w_{49} = 0,495$.

За однакової важливості показників w_8 та w_9 ($B_1 = B_2 = 0,5$) для показника гарантоздатності w_6 (проект Ω_1) обчислимо останній на основі даних, які зведені в табл. 3.

Таблиця 3. Дані для розрахунку показника гарантоздатності першого проекту

Показник	Терми				
	ДНЗ	НЗ	СЗ	ВЗ	ДВЗ
w_8	$\mu_{11} = 0$	$\mu_{12} = 0$	$\mu_{13} = 0$	$\mu_{14} = 1$	$\mu_{15} = 0$
w_9	$\mu_{21} = 0$	$\mu_{22} = 0$	$\mu_{23} = 0,54$	$\mu_{24} = 0,46$	$\mu_{25} = 0$
$x_{j6} = \sum_{i=1}^2 B_i \mu_{ij}$	0	0	0,27	0,73	0

Отже маємо:

$$w_{16} = \sum_{j=1}^5 x_j E_{\alpha_j} = 0 * 0,05 + 0 * 0,25 + 0,27 * 0,45 + 0,73 * 0,65 + 0 * 0,9 = 0,6.$$

При $w_{16} = 0,6$ гарантоздатність проекту Ω_1 , що досліджується, відповідає терму ВЗ з функцією належності $\mu_4(0,6) = 1$, а всім іншим термам — з функцією належності, рівною нулю.

Так само знаходимо показники гарантоздатності для решти трьох проектів. Для проекту Ω_2 відповідно маємо $w_{26} = 0,483$, для проекту Ω_3 — $w_{36} = 0,536$, а для проекту Ω_4 отримали $w_{46} = 0,537$. На основі всіх часткових показників проектів, що зведені в табл. 4, обчислимо оцінки загального показника проектів. Відповідно до порядку ранжування часткових показників коефіцієнти пріоритетності B_k , які отримані за методом Фішберна, дорівнюють:

$$B_1 = \frac{2(7+1-1)}{7*8} = \frac{1}{4}; B_2 = \frac{6}{28}; B_3 = \frac{5}{28}; B_4 = \frac{1}{7}; B_5 = \frac{3}{28}; B_6 = \frac{1}{14}; B_7 = \frac{1}{28}.$$

Таблиця 4. Часткові показники проектів

Проект	Часткові критерії ПАС						
	w_1	w_2	w_3	w_4	w_5	w_6	w_7
Ω_1	ДВЗ Д 0.8_	ДВЗ Д 0.8_	ВЗ Д 1_	ВЗ Д 0.7 і ДВЗ Д 0.3	ВЗ Д 0.4 і ДВЗ Д 0.6_	ВЗ Д 1_	НЗ Д 1_
Ω_2	ВЗ Д 0.9 і ДВЗ Д 0.1	ВЗ Д 0.7 і ДВЗ Д 0.3	ВЗ Д 0.4 і ДВЗ Д 0.6	ВЗ Д 0.7 і ДВЗ Д 0.3	ВЗ Д 0.4 і ДВЗ Д 0.6	СЗ Д 1	НЗ Д 0.2 і СЗ Д 0.8
Ω_3	ДВЗ Д 0.8_	ДВЗ Д 0.8_	ВЗ Д 1_	ДВЗ Д 0.8	ВЗ Д 0.4 і ДВЗ Д 0.6	ВЗ Д 0.36 і СЗ Д 0.64	НЗ Д 0.2 і СЗ Д 0.8.
Ω_4	ВЗ Д 0.9 і ДВЗ Д 0.1	ВЗ Д 0.7 і ДВЗ Д 0.3	ВЗ Д 0.4 і ДВЗ Д 0.6	ДВЗ Д 0.8	ВЗ Д 0.4 і ДВЗ Д 0.6.	ВЗ Д 0.36 і СЗ Д 0.64	СЗ Д 1

Результати розрахунку функцій належності для кожного показника проекту Ω_1 наведені в табл. 5.

Таблиця 5. Функції належності часткових показників першого проекту

Показник	Терми				
	ДНЗ	НЗ	СЗ	ВЗ	ДВЗ
$w_{11} = 0,78$	$\mu_{11}(w_{11}) = 0$	$\mu_{12}(w_{11}) = 0$	$\mu_{13}(w_{11}) = 0$	$\mu_{14}(w_{11}) = 0,2$	$\mu_{15}(w_{11}) = 0,8$
$w_{12} = 0,78$	$\mu_{21}(w_{12}) = 0$	$\mu_{22}(w_{12}) = 0$	$\mu_{23}(w_{12}) = 0$	$\mu_{24}(w_{12}) = 0,2$	$\mu_{25}(w_{12}) = 0,8$
$w_{13} = 0,65$	$\mu_{31}(w_{13}) = 0$	$\mu_{32}(w_{13}) = 0$	$\mu_{33}(w_{13}) = 0$	$\mu_{34}(w_{13}) = 1$	$\mu_{35}(w_{13}) = 0$
$w_{14} = 0,73$	$\mu_{41}(w_{14}) = 0$	$\mu_{42}(w_{14}) = 0$	$\mu_{43}(w_{14}) = 0$	$\mu_{44}(w_{14}) = 0,7$	$\mu_{45}(w_{14}) = 0,3$
$w_{15} = 0,76$	$\mu_{51}(w_{15}) = 0$	$\mu_{52}(w_{15}) = 0$	$\mu_{53}(w_{15}) = 0$	$\mu_{54}(w_{15}) = 0,4$	$\mu_{55}(w_{15}) = 0,6$
$w_{16} = 0,6$	$\mu_{61}(w_{16}) = 0$	$\mu_{62}(w_{16}) = 0$	$\mu_{63}(w_{16}) = 0$	$\mu_{64}(w_{16}) = 1$	$\mu_{65}(w_{16}) = 0$
$w_{17} = 5$	$\mu_{71}(w_{17}) = 0$	$\mu_{72}(w_{17}) = 1$	$\mu_{73}(w_{17}) = 0$	$\mu_{74}(w_{17}) = 0$	$\mu_{75}(w_{17}) = 0$
$x_j = \sum_{i=1}^7 B_i \mu_{ij}$	0	0,036	0	0,486	0,478

Загальний показник для проекту Ω_1 :

$$x_1 = \sum_{j=1}^5 x_j E_{\alpha_j} = 0 * 0,05 + 0,036 * 0,25 + 0 * 0,45 + 0,486 * 0,65 + 0,478 * 0,9 = 0,755.$$

При $x_1 = 0,755$ загальний показник досліджуваного проекту Ω_1 відповідає терму ВЗ з функцією належності $\mu_4(0,755) = 0,45$ та терму ДВЗ з функцією належності $0,55$. У такий же спосіб знаходимо узагальнені показники для трьох інших проектів. Для проекту Ω_2 маємо $x_2 = 0,703$ (відповідає терму ВЗ з функцією належності $\mu_4(0,703) = 0,97$ і терму ДВЗ з функцією належності $\mu_5(0,703) = 0,03$). Для проекту Ω_3 отримали $x_3 = 0,739$ (відповідає терму ВЗ з функцією належності $\mu_4(0,739) = 0,61$ і терму ДВЗ з функцією належності $\mu_5(0,739) = 0,39$). Для проекту Ω_4 отримуємо $x_4 = 0,757$ (відповідає терму ВЗ з функцією належності $\mu_4(0,757) = 0,43$ і терму ДВЗ з функцією належності $\mu_5(0,757) = 0,57$). За узагальненим показником найкращим є проект Ω_4 , найгіршим — проект Ω_2 .

Висновки

Розроблено формалізований підхід до проектування функціонально-орієнтованих інтелектуальних інформаційно-аналітичних систем, який базується на реалізації стандартних функціональних профілів систем даного типу за допомогою попередньо створених, зокрема при виконанні інших проектів, уніфікованих спеціалізованих програмних модулів.

Для оптимізації результату проектування на основі застосування апарату лінгвістичних змінних запропоновано методика аналізу альтернативних проектів інтелектуальної інформаційно-аналітичної системи, котра розробляється, та остаточного вибору кінцевого проекту. Методика на основі лінгвістичних змінних може враховувати суб'єктивність користувача при аналізі та виборі проектів.

1. Жилияев І.Б. Развитие информационного общества в Украине: Правовое регулирование в сфере информационных отношений / И.Б. Жилияев, М.З. Згуровський, М.К. Родіонов. — К.: НТУУ «КПІ». — 2006. — 542 с.
2. Ландэ Д.В. Поиск знаний в Internet. Профессиональная работа / Д.В. Ландэ. — М.: Издательский дом «Вильямс». — 2005. — 272 с.
3. Баранов А.В. DATA MINING. Теория и практика / А.В. Баранов, И.Н. Брянцев, И.М. Жевлаков. — М.: Издательская группа «БДЦ-пресс». — 2006. — 208 с.
4. Барсегян А.А. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP / А.А. Барсегян, М.С. Куприянов, В.В. Степаненко, И.И. Холод. — СПб.: БХВ-Петербург. — 2007. — 384 с.
5. Безсуднов И.В. Интернетика: Навигация в сложных сетях: модели и алгоритмы / И.В. Безсуднов, А.А. Снарский, Д.В. Ландэ. — М.: Либликом. — 2009. — 264 с.
6. Волков И. Архитектура современной информационно-аналитической системы / И. Волков И. Галяхов // Директор информационной службы. — 2002. — № 3.
7. Ларман К. Применение UML 2.0 и шаблонов проектирования. Введение в объектно-ориентированный анализ и проектирование / К. Ларман. — М.: Издательский дом «Вильямс». — 2008. — 736 с.
8. Соммервилл И. Инженерия программного обеспечения. — [6-е изд.] / Соммервилл И. — М.: «Вильямс». — 2002. — 642 с.

9. *Фабрики* разработки программ: потоковая сборка типовых приложений, моделирование, структуры и инструменты / Д. Гринфилд, К. Шорт, С. Кук [и др.]. — М.: «Диалектика». — 2006. — 592 с.
10. *Функциональність* та архітектура сучасних інформаційно-аналітичних систем: Матеріали Міжнародної науково-практичної конференції «Сучасні проблеми і досягнення в галузі радіотехніки, телекомунікацій та інформаційних технологій» / К.С. Бормотова, О.О. Лещенко, А.Ю. Михайлюк [та ін.]. — Запоріжжя. — 2006. — С. 124–125
11. *Кебкало О.С.* Функціональні профілі спеціалізованих інформаційно-аналітичних систем / О.С. Кебкало, А.Ю. Михайлюк, В.П. Тарасенко // Науковий вісник Чернівецького університету: зб. наук. праць: Фізика. Електроніка / Тематичний випуск «Комп'ютерні системи та компоненти». — Частина I. — Чернівці: ЧНУ. — Вип. 423. — С. 117–123.
12. *Критерії* оцінювання захищеності інформації в комп'ютерних системах від несанкціонованого доступу: НД ТЗІ 2.5-004-99 / Затверджено наказом ДСТСЗІ СБ України від 28.04.1999. — № 22.
13. *Класифікація* автоматизованих систем і стандартні функціональні профілі захищеності обробленої інформації від несанкціонованого доступу: НД ТЗІ 2.5-005-99 / Затверджено наказом ДСТСЗІ СБ України від 28.04.1999. — № 22.
14. *Филатов В.А.* Мультиагентные технологии интеграции гетерогенных информационных систем и распределенных баз данных : Дис. д-ра техн. наук: 05.13.06 / В.А. Филатов. — Х: Харьковский национальный ун-т радиоэлектроники. — 2004. — 341 с.: ил. — Библиогр.: С. 313–336.
15. *Заболотня Т.М.* Спеціалізовані інтелектуальні агенти як засіб інтеграції гетерогенного програмного забезпечення / Т.М. Заболотня, В.П. Тарасенко, А.Ю. Михайлик. — Інформаційні технології та комп'ютерна інженерія. — 2006. — № 3(7). — С. 96–101.
16. *Поспелов Б.А.* Нечеткие множества в моделях управления и искусственного интеллекта / Под ред. Б.А. Поспелова. — М.: Наука. — 1986. — 32 с.
17. *Леоненков А.В.* Нечеткое моделирование / А.В. Леоненков. — СПб.: ВНУ. — Петербург, 2003. — 736 с.
18. *Василевич Л.Ф.* Количественные методы принятия решений в условиях риска / Л.Ф. Василевич, К.Н. Маловик, С.Б. Смирнов. — Севастополь: СМУЭИ П. — 2006. — 232 с.
19. *Григорьев В.А.* Оценка качества интеллектуальных систем, оперирующих неформальными понятиями [Электронный ресурс] / В.А. Григорьев // Наука и образование. — 2004. — № 10. — Режим доступа: URL:<http://technomag.edu.ru/sdocs/about.html>
20. *Лахутин Д.Г.* О проблеме оценки поисковых систем. Ч. 1. / Д.Г. Лахутин, С.О. Чернявский // НТИ. — 1970. — Сер. 2. — № 1. — С. 24–34.
21. Экспертные оценки в научно-техническом прогнозировании / [Г.М. Добров, Ю.В. Ершов, Е.И. Левин, Л.П. Смирнов]. — К.: Наук. думка, 1974. — 160 с.

Надійшла до редакції 22.06.2010