

# Формалізація Етапів Діагностично-Лікувальних Заходів при Проектуванні Систем Підтримки Прийняття Рішень в Медицині

Анатолій Поворознюк  
кафедра обчислювальної техніки та програмування  
Національний технічний університет  
«Харківський політехнічний інститут»  
Харків, Україна  
ai.povoroznjuk@gmail.com

Ганна Філатова  
кафедра обчислювальної техніки та програмування  
Національний технічний університет  
«Харківський політехнічний інститут»  
Харків, Україна  
filatova@gmail.com

## Formalization of the Stages of Diagnostic and Treatment Activities in the Design of Decision Support Systems in Medicine

Anatoliy Povoroznjuk  
Department of Computer Science and Programming  
National Technical University  
"Kharkiv Polytechnic Institute"  
Kharkiv, Ukraine  
ai.povoroznjuk@gmail.com

Anna Filatova  
Department of Computer Science and Programming  
National Technical University  
"Kharkiv Polytechnic Institute"  
Kharkiv, Ukraine  
filatova@gmail.com

**Анотація**—Формалізовано етапи діагностично-лікувального процесу при проектуванні комп'ютерних систем підтримки прийняття рішень в медицині. Розроблено математичну модель процесу діагностики та лікарських дій з метою підвищення ефективності надання медичних послуг та мінімізації ризиків лікарських помилок.

**Abstract**—The stages of the diagnostic and therapeutic process in the design of computer decision support systems in medicine are formalized. The mathematical model of the process of diagnostics and medical actions for the purpose of increasing the efficiency of providing medical services and minimizing the risks of medical errors is developed.

**Ключові слова**—комп'ютерна система; діагностика; лікарська дія; математична модель; лікарська помилка.

**Keywords**— computer system, diagnostics, medical action, mathematical model, medical error.

### I. ВСТУП

Комплекс діагностично-лікувальних заходів (ДЛЗ) складається з двох взаємопов'язаних етапів: діагностики захворювань і лікування виявлених патологій, причому після діагностики і призначення лікувальних процедур

необхідний моніторинг поточного стану пацієнта з метою оцінки ефективності процесу лікування і, при необхідності, його корекції. Для лікування того чи іншого захворювання необхідно надання певних лікарських дій (ЛД) на організм (хірургічне втручання, фармакологічний, лікувально-терапевтичний вплив, реабілітаційні заходи).

На кожному із зазначених етапів лікар, як особа, яка приймає рішення, виробляє управлінське рішення в умовах дефіциту вихідних даних і суттєвої апріорної невизначеності, ґрунтуючись на своїй кваліфікації, досвіді і інтуїції. При цьому прийняття неправильного рішення (лікарська помилка) як на етапі діагностики, так і на етапі лікування може мати катастрофічні наслідки для здоров'я пацієнта. Термін «лікарська помилка» (ЛП) визначає неправильну діагностику хвороби або неправильні ЛД, які обумовлені добросовісною помилкою лікаря, при цьому виключається недбалість і несумлінність при виконанні своїх обов'язків. Причиною помилки діагностики є недостатній обсяг діагностичних даних, або їх невірна інтерпретація (особливо при суб'єктивному аналізі якісних показників). Причиною неправильних ЛД при правильному діагнозі є недостатнє врахування індивідуальних особливостей пацієнта (алергічні реакції на

певні препарати, список хвороб, якими вже хворів пацієнт, які ліки приймав і т.д.).

В даний час є широкий спектр комп'ютерних діагностичних систем в різних предметних областях медицини [1, 2], інформатизація ЛД обмежується медичними довідниками, в тому числі у вигляді інформаційно-пошукових систем [3]. В сучасних комп'ютерних системах підтримки прийняття рішень в медицині задачі діагностики і ЛД розглядаються незалежно один від одного, тому актуальною є задача мінімізації ризику ЛП при комплексній оцінці діагностичних і лікарських дій (ДЛД).

Метою роботи є формалізація етапів та розробка математичної моделі ДЛД з метою підвищення їх ефективності та мінімізації ризику лікарських помилок.

## II. РОЗРОБКА МАТЕМАТИЧНОЇ МОДЕЛІ ДІАГНОСТИЧНО-ЛІКУВАЛЬНОГО ПРОЦЕСУ

На основі аналізу етапів ДЛД [2] пропонується наступна математична модель  $M_{da}$ :

$$M_{da} = \langle X, S, B, \hat{S}, \hat{B}, D, M_a, \tilde{D}, f, Q_f \rangle,$$

де  $X = \{x_i(\cdot) | i \in \{1, 2, \dots, n_x\}\}$  – множина реєструємих біомедичних сигналів/зображень БМС/З ( $n_x$  – кількість БМС/З). Так як БМС є деякою функцією одного аргументу, а БМЗ – двох аргументів, то для уніфікації розгляду процесу обробки біомедицинських сигналів і зображень аргументи функції, що задає сигнал або зображення, вилучимо, тобто  $i$ -й БМС/З позначимо  $x_i(\cdot)$ ;

$S = \{s_j(\cdot) | j \in \{1, 2, \dots, n_s\}\}$  – множина діагностичних ознак ( $n_s$  – кількість діагностичних ознак);

$B$  – множина можливих комбінацій діапазонів значень діагностичних ознак  $S$ ;

$\hat{S} = \langle \hat{s}_1, \dots, \hat{s}_{n_s} \rangle$  – упорядкована множина значень діагностичних ознак, отриманих для конкретного пацієнта, включаючи інструментальне обстеження;

$\hat{B} = \langle \hat{b}_1, \dots, \hat{b}_{n_s} \rangle$  – упорядкована множина значень діапазонів для  $\hat{S}$ ;

$D = \{d_k | k \in \{1, 2, \dots, n_D\}\}$  – множина можливих діагнозів у заданій предметній області, тобто алфавіт діагнозів ( $n_D$  – кількість діагнозів);

$M_a = \{m_{a_k} | k \in \{1, 2, \dots, n_{M_a}\}\}$  – множина лікарських дій ( $n_{M_a}$  – кількість лікарських дій);

$\tilde{D} = \{\tilde{d}_j | j \in \{1, 2, \dots, n_D\}\}$  – нечітка множина діагностичних висновків;

$f$  – множина відображень відповідних множин  $X, S, B, \hat{B}, D, \hat{M}_a$ ;

$Q_f$  – множина критеріїв відображень  $f$ ,  $n_x, n_s, n_D, n_{M_a} \in N$ ;  $N$  – множина натуральних чисел.

Розглянемо більш детально елементи запропонованої математичної моделі  $M_{da}$ , які формалізують відповідні етапи ДЛЗ. У процесі інструментального обстеження реєструється множина БМС/З  $X = \{x_i(\cdot) | i \in \{1, 2, \dots, n_x\}\}$ , передбачених протоколом обстеження в даній предметній області. Наприклад, у результаті електрокардіографічного обстеження формується множина відведень ЕКГ, у результаті мамографічного обстеження – множина проєкцій лівої й правої молочних залоз і так далі.

В результаті обробки кожного БМС/З  $x_i(\cdot)$  визначається підмножина діагностичних ознак  $S_i = \{s_{ij} | j = \overline{1, n_i}\}$ , де  $n_i$  – кількість діагностичних ознак для  $i$ -го БМС/З, причому  $S_i \cap S_j = \emptyset$ , якщо  $i \neq j$ , і  $\bigcup_{i=1}^{n_x} S_i = S$  – множина діагностичних ознак для всіх БМС/З, тобто для множини  $X$ .

Для кожної з діагностичних ознак  $s_j$  ( $j = \overline{1, n_s}$ ) визначені діапазони значень  $B_j = \{b_{jk} | k = \overline{1, n_{B_j}}\}$ , де  $n_{B_j}$  – кількість діапазонів значень для діагностичної ознаки  $s_j$ . Розбивка припустимого інтервалу значень числової ознаки  $s_j$  на діапазони  $b_{jk}$  відповідає значенням типу «норма», «нижче/вище норми», «граничні значення», «небезпечні значення», «критичні значення» і так далі, прийнятим у заданій предметній області медицини. Число діапазонів  $n_{B_j}$  може бути різним у різних ознак.

Множина можливих комбінацій діапазонів діагностичних ознак це прямий (декартовий) добуток множин  $B_j$ , тобто  $B = B_1 \times \dots \times B_j \times \dots \times B_{n_s} = \{b_1, \dots, b_j, \dots, b_{n_s} | b_j \in B_j, j \in \{1, 2, \dots, n_s\}\}$ .

В результаті реєстрації й аналізу БМС/З формується кортеж значень діагностичних ознак  $\hat{S} = \langle \hat{s}_1, \dots, \hat{s}_{n_s} \rangle$  для конкретного пацієнта. Тоді кортеж оцінок відповідних діапазонів  $\hat{B} = \langle \hat{b}_1, \dots, \hat{b}_{n_s} \rangle \in B$ , це реалізація  $B$  для даного пацієнта, тобто представлення результатів вимірів точкою в просторі  $B$ .

Проєкції  $B$  на можливі комбінації координатних осей утворюють множину можливих симптомокомплексів  $C = \{pr_i B, pr_{i_1 \dots i_j \dots i_q} B | i, i_j, q \in \{1, \dots, n_s\}\}$ . Із цієї множини виділяється підмножина інформативних симптомокомплексів  $C' = \{c'_i : i = \overline{1, n_{C'}}\} \subseteq C$  ( $n_{C'}$  – кількість інформативних симптомокомплексів) щодо системи діагнозів  $D$ .

Проекції  $\hat{B}$  на можливі комбінації координатних осей утворюють множину можливих кортежів, які складаються із можливих комбінацій діагностичних ознак

$$\hat{C} = \{pr_i \hat{B}, pr_{i_1, \dots, i_q} \hat{B} \mid i, i_j, q \in \{1, \dots, n_S\}\} = \{\hat{c}_i \mid i = \overline{1, n_{\hat{C}}}\},$$

де  $n_{\hat{C}}$  – кількість комбінацій діагностичних ознак.

Тоді множину інформативних комбінацій ознак щодо системи діагнозів  $D$  можна визначити як

$$\hat{C}' = \{\hat{c}_i \mid \hat{c}_i \in C'_i, i = \overline{1, n_{C'}}, j = \overline{1, n_{C'}}\} = \{\hat{c}'_i \mid i = \overline{1, n_{\hat{C}'}}\}.$$

Для реалізації системи підтримки прийняття рішень в медицині (СППРМ) необхідною умовою є те, що множина можливих діагнозів  $D = \{d_k \mid k \in \{1, 2, \dots, n_D\}\}$ , повинна утворювати повну групу несумісних подій, тобто

$$P\left(\bigcup_{k=1}^{n_D} d_k\right) = \sum_{k=1}^{n_D} P(d_k) = 1.$$

Нехай множина  $D = \{d_k \mid k \in \{1, 2, \dots, n_D\}\}$  є універсальною множиною діагнозів. Нехай задано сімейство характеристичних функцій  $\mu_j(d_k)$ ,  $j = \overline{1, n_{\hat{D}}}$ , що показують належність  $k$ -го діагнозу  $j$ -му діагностичному висновку. Тоді нечітка підмножина  $\tilde{D}_j = \{(d_k, \mu_j(d_k)) \mid d_k \in D\}$  відповідає нечіткому поняттю « $j$ -й діагностичний висновок». Позначимо множину нечітких понять як  $\tilde{D} = \{\tilde{D}_j \mid j \in \{1, \dots, n_{\tilde{D}}\}\}$ , де  $n_{\tilde{D}}$  – число діагностичних висновків.

Характеристична функція  $\mu_j(d_k)$  приймає значення з лінійно упорядкованої множини належностей  $M$ . Якщо для діагностики застосовується детерміністична логіка, то  $M = \{0, 1\}$ , і нечітка множина розглядається як чітка, причому  $\tilde{D}_j = \{d_k \mid \mu_j(d_k) = 1\}$ .

При застосуванні імовірнісної логіки кожний діагноз  $d_k$  задається умовною ймовірністю  $p(d_k \mid \hat{c}'_j)$ , яка виступає в якості характеристичних функцій  $\mu_j(d_k)$ , тобто  $\tilde{D}_j = \{(d_k, \mu_j(d_k)) \mid d_k \in D, \mu_j(d_k) = p(d_k \mid \hat{c}'_j)\}$ .

Крім розглянутих вище множин у математичній моделі  $M_{da}$  задається множина відображень  $f = \{f_{X\hat{S}}, f_{\hat{S}\hat{B}}, f_{\hat{B}\hat{C}'}, f_{\hat{C}'\tilde{D}}, f_{\tilde{D}M_a}\}$  і множина критеріїв  $Q = \{Q_{X\hat{S}}, Q_{\hat{B}\hat{C}'}, Q_{\hat{C}'\tilde{D}}, Q_{\tilde{D}M_a}\}$  для цих відображень. Розглянемо докладніше компоненти цих множин.

Функція  $f_{X\hat{S}}$  у відповідність із підмножиною критеріїв  $Q_{X\hat{S}}$  і множиною діагностичних ознак  $S$  задає відображення множин БМС/З  $X$  на множину значень діагностичних ознак  $\hat{S}$ , тобто  $f_{X\hat{S}} : X \rightarrow \hat{S}$ . Дане відображення реалізується процедурою морфологічного аналізу БМС/З [4,5].

Функція  $f_{\hat{S}\hat{B}}$  задає бієктивне відображення множини значень діагностичних ознак  $\hat{S}$  на множину оцінок діапазонів діагностичних ознак  $\hat{B}$ , тобто  $f_{\hat{S}\hat{B}} : \hat{S} \rightarrow \hat{B}$ .

Функція  $f_{\hat{B}\hat{C}'}$  у відповідність із підмножиною критеріїв  $Q_{\hat{B}\hat{C}'}$  задає відображення множини оцінок діапазонів діагностичних ознак  $\hat{B}$  на множину оцінок симптомокомплексів  $\hat{C}'$ , тобто  $f_{\hat{B}\hat{C}'} : \hat{B} \rightarrow \hat{C}'$ .

Інформативність (корисність) ознак, що входять у симптомокомплекс, визначається експертними оцінками або розрахунками інформативності ознак по навчальній вибірці. Причому симптомокомплекси повинні мати не тільки інформативність, але й повнотою опису, компактністю й несуперечністю ознак, що входять у нього.

Функція  $f_{\hat{C}'\tilde{D}}$  у відповідність із підмножиною критеріїв  $Q_{\hat{C}'\tilde{D}}$  задає відображення множини оцінок симптомокомплексів  $\hat{C}'$  на множину діагностичних висновків, тобто  $f_{\hat{C}'\tilde{D}} : \hat{C}' \rightarrow \tilde{D}$ . При цьому виконується синтез ієрархічної структури діагностуємих станів в заданій предметній області медицини (бінарне дерево рішень  $D_S$ ), процедурою ієрархічної кластеризації множини діагностуємих станів  $\{D_i\}_{n_D}$  за критерієм мінімуму помилки кластеризації в просторі ознак  $S$ .

Результатом кластеризації є бінарне дерево рішень  $D_S$ , коренем якого є повна множина діагнозів  $\{D_i\}_{n_D}$  в заданій предметній області, в гілках розташовуються кластери діагнозів, а листками є окремі діагнози. В процесі діагностики  $i$ -го пацієнта при відомому векторі ознак  $\hat{S}_i$  відбувається рух по дереву рішень, в кожній  $j$ -й вершині якого виконується диференційна діагностика станів  $D_q$  та  $D_l$ , шляхом обчислення вирішального правила (ВП) і прийняття рішення на користь  $D_q$  або  $D_l$ .

Для реалізації імовірнісного ВП в роботі реалізується метод синтезу уточнюючого діагнозу [2], який є модифікацією методу послідовного аналізу (методу Вальда). На кожному  $k$ -му етапі ВП ( $k = \overline{1, n_k}$ ), при диференційній діагностиці між двома станами  $D_q$  і  $D_l$ , аналізується чергова ознака  $\hat{s}_k$  і вираховується відношення правдоподібності

$$\Theta = \prod_i P(\hat{b}_k / D_q) / P(\hat{b}_k / D_l),$$

яке порівняється з порогоми  $\Theta > A$ ,  $\Theta < B$ .

При виконанні однієї з умов приймається рішення про  $D_q$  або  $D_l$  відповідно й виконується перехід на більш низький рівень ієрархії  $D_S$  з метою уточнення діагнозу. При невиконанні обох нерівностей додається наступна  $i + 1$  ознака й процедура повторюється.

Ризики неправильного прийняття рішення на етапі діагностики:  $\alpha$  – помилка першого роду і  $\beta$  – помилка

другого роду, визначаються розташуванням еліпсоїдів розсіювання об'єктів навчальної вибірки в просторі ознак  $S$  без урахування їх впливу на етап вибору необхідних ЛД і їх подальшої реалізації.

Функція  $f_{\tilde{D}_{M_a}}$  у відповідність із критерієм  $Q_{\tilde{D}_{M_a}}$  задає відображення множини діагностичних висновків  $\tilde{D}$  на множину лікарських дій  $M_a$ , тобто  $f_{\tilde{D}_{M_a}}: \tilde{D} \rightarrow M_a$ .

ЛД представляються моделлю  $M_a = \langle Ta, Pa, SI \rangle$ , де  $M_a$  – множина ЛД;  $Ta = \{ta_i | i \in \{1, 2, \dots, n_{ta}\}\}$  – множина терапевтичних дій (ТД),  $Pa = \{pa_i | i \in \{1, 2, \dots, n_{pa}\}\}$  – множина фармакологічних дій (ФД),  $SI = \{si_i | i \in \{1, 2, \dots, n_{si}\}\}$  – множина видів хірургічного втручання.

Призначення ЛД при відомому діагностичному висновку  $\tilde{D}_i$  складається з визначення їх типу ( $Ta$ ,  $Pa$ ,  $SI$ , або їх комбінацій) і переліку конкретних дій. Вибір типу ЛД є задачею багатокритеріального вибору альтернатив, для вирішення якої використовується метод аналізу ієрархій (МАІ). Для кожного з допустимих для даного діагностичного висновку  $\tilde{D}_i$  типу ЛД формується підмножина необхідних ЛД  $ta_{D_i} \in Ta$ ,  $pa_{D_i} \in Pa$ ,  $si_{D_i} \in SI$ , після чого визначається їх реалізація з урахуванням індивідуальних особливостей пацієнта, протипоказань до окремих ЛД і багатокритеріального вибору аналогів.

Для комплексної оцінки етапів ДЛД і мінімізації ризику ЛПІ шукається залежність між помилкою діагностики ( $D_q$  замість  $D_i$ ) і її наслідків при реалізації ЛД. Так як для переважної кількості патологій в різних областях медицини ЛД реалізуються медикаментозним шляхом, розглянемо такий варіант реалізації ЛД, при якому кожному діагностичному висновку  $\tilde{D}_i$  відповідає множина необхідних ФД  $pa_{D_i}$ , на підставі якого формується комплекс лікарських препаратів (КЛП)  $Y_i$ , який забезпечує реалізацію  $pa_{D_i}$ , з урахуванням непереносимості  $i$ -го пацієнта до окремих препаратів і багатокритеріального порівняння препаратів-аналогів. Так як  $Y_q \rightarrow pa_{D_q}$ , а  $Y_i \rightarrow pa_{D_i}$ , то ризик ЛПІ визначається розбіжністю компонентів множин  $pa_{D_q}$  та  $pa_{D_i}$ , і для його мінімізації виконується перехід від традиційного простору ознак  $S$  в простір ФД  $Pa$ , компонентами якого є бінарні змінні (0 – відсутня ФД, 1 – присутня), а кожен стан  $\tilde{D}_i$  представляється  $i$ -ю вершиною гіперкуба. В якості міри близькості в просторі  $Pa$  вибрана зважена відстань Хеммінга:

$$r_{ij} = \sum_{k=1}^g w_{ij} |pa_{ki} - pa_{kj}|,$$

де  $pa_{ki}, pa_{kj} \in \{0, 1\}$  –  $k$ -та ФД  $i$ -го та  $j$ -го діагнозів відповідно;  $g$  – розмірність простору  $Pa$ ;  $w_{ij}$  – коефіцієнт, який забезпечує збільшення відстані, в разі наявності конфліктуючих ФД.

Застосування ієрархічної кластеризації за критерієм мінімуму сумарної зв'язку (мінімальний розріз  $R$ ) в просторі  $Pa$  забезпечує мінімум ризику прийняття рішення на етапі формування КЛП при синтезі дерева рішень  $D_S$  на етапі діагностики. Крім того, в роботі пропонується метод корекції порогів  $A$  і  $B$  в ВП, враховуючи помилки, які виникають на етапі призначення КЛП. Отримано залежності між  $\alpha$ ,  $\beta$  і нормованим значенням мінімального розрізу  $\bar{R}_i$ :  $\alpha_n = 0,5(1 - \bar{R}_{q_i})$ ,  $\beta_n = 0,5(1 - \bar{R}_{i_q})$ . Визначені таким чином  $\alpha$  і  $\beta$  задають пороги  $A = (1 - \beta) / \alpha$ ,  $A = \beta / (1 - \alpha)$  в ВП, що забезпечує врахування ризиків ЛПІ при призначенні КЛП в діагностичному ВП.

Розглянута технологія комплексної оцінки ДЛД з метою мінімізації ризиків ЛПІ адаптується до інших видів ЛД, при цьому виконується перехід з простору ознак  $S$  у простір відповідних ЛД ( $Ta$  або  $SI$ ).

Виконана програмна реалізація системи і її тестування на реальних медичних даних з використанням навчальної вибірки з 400 пацієнтів. В якості основної платформи обґрунтований вибір Java. В архітектурі системи виділено три основних модуля: модуль взаємодії з користувачем, базу даних, що включає в себе базу знань і модуль побудови знань.

### III. ВИСНОВКИ

Виконано формалізацію етапів проведення ДЛД. Розроблено математичну модель ДЛД, яка враховує комплексну оцінку її складових, що дозволяє мінімізувати ризики лікарських помилок, підвищити достовірність і обґрунтованість рішень. Архітектура програмного забезпечення системи дозволяє легко адаптуватися до різних предметних областей медицини.

### ЛІТЕРАТУРА REFERENCES

- [1] Об инфраструктуре информационной поддержки клинической медицины / В.А. Лишук, А.В. Гаврилов, Г.В. Шевченко и др. // Медицинская техника. – М.: 2003. – № 4. – С. 36-42.
- [2] Поворознюк А.И. Системы поддержки принятия решений в медицинской диагностике. Синтез структурированных моделей и решающих правил / А.И. Поворознюк – Saarbrücken Germany: LAP LAMBERT Academic Publishing GmbH & Co. KG, 2011. – 314 с.
- [3] Компендиум 2015 – лекарственные препараты /Под ред. В.Н. Коваленко, А.П. Викторова. [Электронный ресурс] <http://www.compendium.com.ua>.
- [4] Пытьев Ю.П. Косые проекторы и относительные формы в морфологии / Ю.П. Пытьев // Журнал вычислительной математики и математической физики. – 2013. – № 53 (12). – С. 2100-2122
- [5] Филатова А.Е. Морфологическая фильтрация полутоновых изображений на основе локальных статистик / А.Е. Филатова // Автомат. технологии и производства. – 2016. – №3(13). – С. 33-39.