

УДК 004

Ю.Гордієнко, магістр гр. КІ-22МЗ,
Центральноукраїнський національний технічний університет

ДОСЛІДЖЕННЯ ТА ПРОГРАМНА РЕАЛІЗАЦІЯ СИСТЕМИ ІДЕНТИФІКАЦІЇ ПРОФЕСІЙНИХ ЗНАНЬ ОПЕРАТОРІВ АВТОМАТИЗОВАНИХ СИСТЕМ УПРАВЛІННЯ

У статті розроблено програмне забезпечення, яке призначено для системи ідентифікації професійних знань операторів автоматизованих систем управління. Метою розробки є дослідження та програмна реалізація системи ідентифікації професійних знань операторів автоматизованих систем управління. Об'єктом дослідження є процес ідентифікації професійних знань операторів автоматизованих систем управління. Предметом дослідження є методи ідентифікації професійних знань операторів автоматизованих систем управління. Методи дослідження базуються на методах теорії комп'ютерного тестування, методах математичної статистики, методах розробки програмного забезпечення. Результат роботи – програмна реалізація системи ідентифікації професійних знань операторів автоматизованих систем управління. В процесі роботи над програмною моделлю виконано аналіз існуючих апаратних та програмних засобів. В повній мірі описані всі компоненти розробленого програмного забезпечення.

Постановка проблеми. Визнаним елементом промислової безпеки є комп'ютерне навчання операторів процесів. Значна кількість нещасних випадків, демографічні проблеми та брак висококваліфікованого персоналу автоматизованих систем управління вимагають ефективних методів навчання. Досягнення в області систем навчання операторів (СНО АСУ) є значними: ринок СНО АСУ величезний (1 мільярд доларів на рік, включаючи відповідні високотехнологічні питання промислової автоматизації) з усіма СНО АСУ. компоненти є високорозвиненими (математичні моделі та середовища керування є складними та точними, інструменти інструкцій є досить розвиненими). СНО АСУ хороші для інвестицій у промислову безпеку. Середній ефект від використання однієї системи навчання протягом життєвого циклу оцінюється в 18 мільйонів доларів, а річний ефект, розрахований на основі довгострокової статистики аварій з вини оператора, досягає 0,5 долара за тону сирової нафти на рік. Однак, незважаючи на ці досягнення, розробники та користувачі СНО АСУ відчувають нестачу її функціональності. Багато публікацій присвячено досвіду розробки та використання СНО АСУ, це сфера жвавих дискусій.

Зупинимося на техніці комп'ютерного навчання. Його вплив на якість навчання досліджено не так глибоко, як аспекти моделювання. У цьому контексті дослідницькі зусилля зосереджені на наступному:

- класифікації комп'ютерних систем навчання;
- структурі та кількості навчальних вправ;
- оцінці результатів навчання.

Але питання «Як виміряти вплив навчання на операторів?» Оцінити якість виконання вправ і посттренувальних тестів просто. За результатами можна формально оцінити «фактичний» рівень компетентності оператора. Але ця оцінка не дає відповіді на поставлене вище питання, оскільки вона не гарантує успіху оператора в іншій ситуації. Важливими питаннями є: наскільки успішно навички, отримані в процесі навчання, будуть перенесені в реальну практику, і чи зміниться структура знань оператора в результаті навчання. Окрім цього цікавить залежність успіху оператора від особистих особливостей. Усі ці питання важливі для людей, які відповідають за навчання персоналу. Їм потрібен надійний зворотний

зв'язок від системи навчання, який би відображав об'єктивний рівень компетентності оператора, крім дійсного навчання.

Дослідження в області оцінки структури знань і її характеристик досить рідкісні. Решта цієї статті структурована наступним чином. У наведеному нижче розділі описані різні підходи до покращення якості підготовки операторів. А саме оцінювання змін попередньої підготовки та структури КМ (структури знань та кластерної техніки). У наступному розділі представлено експериментальне дослідження оцінки еволюції СМ, наступний розділ містить обговорення результатів і плани майбутніх тренувань. Останній розділ завершує цю роботу.

Аналіз останніх досліджень і публікацій. При аналізі останніх досліджень і публікацій [1-10] було виявлено певні прогалини у забезпеченні системи ідентифікації професійних знань операторів автоматизованих систем управління.

Мета й завдання дослідження. Метою роботи є дослідження та програмна реалізація системи ідентифікації професійних знань операторів автоматизованих систем управління.

Для досягнення поставленої мети визначена програма дослідження, що складається з наступних завдань:

- Огляд існуючих систем ідентифікації професійних знань операторів автоматизованих систем управління.
- Дослідження системи ідентифікації професійних знань операторів автоматизованих систем управління.
- Програмна реалізація системи ідентифікації професійних знань операторів автоматизованих систем управління.

Об'єктом дослідження є процес ідентифікації професійних знань операторів автоматизованих систем управління.

Предметом дослідження є методи ідентифікації професійних знань операторів автоматизованих систем управління.

Методи дослідження базуються на методах теорії комп'ютерного тестування, методах математичної статистики, методах розробки програмного забезпечення.

Виклад основного матеріалу.

Огляд оцінки результатів навчання

Основним принципом оцінки якості СНО АСУ є відповідність діяльності оператора на навчанні реальній активності. Сучасні системи навчання дозволяють досягти високого рівня такого роду подібності. Вони точні в моделюванні процесів, і середовище керування може бути ідеально змодельоване. Методи навчання базуються на сучасних уявленнях про те, що потрібно розумінню оператора під час керування процесом. Останній включає ієрархічну структуру різних навичок і розумового фону.

Високоякісна система навчання є необхідною умовою успіху оператора, але не гарантією. Багато що залежить від організації навчання. Потрібно вміти оцінювати прогрес стажера і на його основі змінювати програму навчання. Проблема коригування програми підготовки зазначена в регламенті СНО АСУ. Наприклад, він може забезпечити повторне навчання у випадку прогалин у деяких розділах програми. Базуючись на результатах формального навчання або якості реального контролю процесу, його можуть регулювати інструктори з навчання або лінійні керівники слухачів. Відомі підходи до оцінювання професійних навичок персоналу за допомогою структури «обов'язки – компетенції – поведінка – рівні кваліфікації».

Обов'язки оператора можна описати компетенціями, які визначаються професійною роллю (наприклад, керування процесом у звичайному режимі, прогнозування та реагування на наслідки ненормальних ситуацій). Компетенції охоплюють певну поведінку оператора. Наприклад, щоб контролювати процес у нормальному стані, оператор консолі повинен зчитувати передавачі, працювати з елементами керування, взаємодіяти з іншими операторами тощо. Останнє можна виміряти за допомогою сценаріїв навчання, тестів тощо, пов'язаних із поведінкою. Рівень кваліфікації (від «початківця» до «експерта»)

призначається з огляду на вимірювання продуктивності вправ. Ці оцінки важливі для навчання оператора, але вони не показують зміни в когнітивних характеристиках слухача як результат навчання. Рідкісним винятком є публікація, яка поставила питання про те, як навчання може найкраще підтримувати ефективність контролю процесу з урахуванням індивідуальних відмінностей. Зокрема, однією з цілей цієї роботи є показати зв'язок між структурою знань і результатом навчання в управлінні процесом. Підхід, представлений у нашій роботі, базується на дослідженнях структурних характеристик концептуальної моделі оператора.

Про попередню підготовку оператора

Як знання оператора розвиваються під час професійного зростання? Після комплексу початкових тренінгів здається іспит для допуску до роботи та закріплюється навички надалі в трудовій практиці. На якомусь етапі можна зіткнутися з СНО АСУ. Однак для невідповідної людини комплексне комп'ютерне навчання безглузде. По-перше, слід розвинути базові навички контролю. Перш за все, розвинути розуміння причинно-наслідкових зв'язків в системі, що перебуває під керуванням, і навчитися діагностувати причини системних збоїв. Існують технічні засоби та методи попереднього навчання, які були впроваджені в різних заводських навчальних центрах. Нижче наведено огляд комп'ютерної навчальної системи, яка була використана в одному з експериментальних досліджень. Учасник грає в покрокову гру, яка імітує процес прийняття діагностичних рішень у реальній практиці. Гра базується на підмножині збоїв процесу та вимірюваних параметрах стану процесу.

Система випадковим чином вибирає одну з помилок і зберігає її в таємниці. Стажеру повідомляють про один із симптомів (відхилення параметрів від нормальних значень), пов'язаних з невідомою ще несправністю. Завдання стажиста – виявити обрану невдачу. Можна або вибрати одну помилку зі списку та підтвердити її (ця дія «коштує» великого штрафу у разі помилки), або перевірити стан інших параметрів процесу (цей вибір «вартує» мінімальний штраф).

Мета стажера – виявити приховану помилку, отримавши мінімальний загальний штраф. На початку гри багато помилок можуть мати відношення до початкової картини ненормального стану процесу. У цьому випадку вигідніше запросити додаткову інформацію про симптоми, щоб звузити список кандидатів. Ключове питання на цьому етапі – «який параметр краще запитувати?», виходячи з попереднього розуміння наслідків кожної (списку кандидатів) несправності з точки зору відхилення стану процесу. Правильна відповідь на це запитання веде до швидшої стратегії розв'язання. Кінцевий бал розраховується на основі порівняння отриманих штрафних балів і балів, які можна було б отримати, якби слухач перевіряв усі симптоми параметрів процесу. Зауважте, що тривалість розв'язання гри не впливає на результат. Стрес від дефіциту часу на вирішення проблем у реальному житті змінюється нестачею інформації в навчальній грі

З метою розробки методів структурної ідентифікації математичної моделі навчання операторів АСУ розглянемо проблеми ефективності системи професійного навчання операторів АСУ.

Побудову моделі ідентифікації процесу навчання засновано на уявленні останнього процесу навчання у вигляді цілеспрямованого функціонування деякої «складної системи».

Виділено в структурі системи навчання три основні підсистеми: $\{S_v\}$, $v=1, N_s$; операторів $\{P_i\}$ загальною чисельністю N_s ; $\{P_i\}$, $i=1, N_p$, навчальних предметів загальною чисельністю N_p ; $\{O_j\}$, $j=1, N_o$, способів організації навчання загальною чисельністю N_o . Ентропію системи навчання позначено виразом:

$$H_{CO} = -\varphi \cdot \log_2 \varphi - (1 - \varphi) \cdot \log_2 (1 - \varphi) \quad (1)$$

Введені поняття моделі ефективності дозволяють трансформувати постановку задачі побудови моделі навчання, системи навчання, тобто пошуку функціональної

залежності $\dot{Y}_{CO} = f(\dot{Y}_S; \dot{Y}_I; \dot{Y}_W)$. Іншими словами, для одного періоду розгляду діяльності системи навчання або для одного такту навчання можна розглядати ймовірність результатів цього такту у вигляді:

$$\begin{cases} \varphi^{(l)} = \varphi^{(0)} \cdot p_{11}^{(l)} + (1 - \varphi^{(0)}) \cdot p_{01}^{(l)}; \\ 1 - \varphi^{(l)} = \varphi^{(0)} \cdot p_{10}^{(l)} + (1 - \varphi^{(0)}) \cdot p_{00}^{(l)}; \end{cases} \quad (2)$$

ефективність системи навчання визначається через ймовірність успіху при нескінченному числі тактів навчання:

$$\dot{Y}_{CO} = \dot{Y}_{CO}^{(\infty)} = \ln \frac{\varphi^{(\infty)}}{1 - \varphi^{(\infty)}}$$

Безумовна ймовірність успіху системи навчання [CH], тобто рівняння моделі представлено у вигляді:

$$\varphi = \frac{p}{p+q} = \frac{z \cdot \lambda \cdot \gamma}{z \cdot \lambda \cdot \gamma + (1-z) \cdot (1-\lambda) \cdot (1-\gamma)}. \quad (3)$$

Якщо тепер знайти математичну модель ефективності системи

$$\begin{aligned} \dot{Y}_{CO} &= \ln \frac{\varphi}{1-\varphi} = \ln \frac{p/(q+p)}{q/(q+p)} = \ln \frac{p}{q} = \\ &= \ln \frac{z \cdot \lambda \cdot \gamma}{(1-z) \cdot (1-\lambda) \cdot (1-\gamma)} = \dot{Y}_S + \dot{Y}_I + \dot{Y}_\gamma, \end{aligned} \quad (4)$$

то ефективність усієї системи навчання буде визначатися сумою ефективності її елементів. Встановлено, що необхідне врахування ще однієї властивості системи навчання, її автономності, обумовленої коефіцієнтом регресії рангу успішності після будь-якого такту навчання щодо попереднього рангу $W = 1 - p - q$. Автономність системи навчання може бути оцінена через параметри елементів як навчальної програми Π_i так і для всієї кількості респондентів $\{S_v\}$:

$$W = z \cdot (1 - \lambda) + \lambda \cdot (1 - \gamma) + \gamma \cdot (1 - z); \lambda = p(\Pi_i); z = p(S_v); \gamma = p(O_j).$$

У цьому випадку ми можемо стверджувати, що число випадків спостережень $R_{vi} = 1$ (або $R_{vi} = 0$) навіть за один такт функціонування системи навчання несе деяку інформацію про властивості системи навчання і її елементів.

Описано новий метод аналізу первинної інформації на основі її згортки. Для кожного респондента, S_v , $v=1, N_s$, передбачається одна з трьох категорій успішності: N, U, K . Доведено теорему про аперіодичний розподіл балу тесту.

Доведення асимптотичної нормальності розподілу балу респондента цінний ще і тому, що він може бути поширеним і на слабко корельовані тести; функція розподілу нормованої суми слабко зв'язаних випадкових величин із зростом числа доданків прямує до нормального закону. Доведено, що мірою розсіювання бала тесту, тобто вихідного його параметра, є сума математичного очікування міри розсіювання, зумовленого властивостями респондентів. На підставі цього висновку у роботі стверджується, що вірогідний розподіл властивостей респондентів у групі N є неминучим. Тільки лише у випадку «детермінованого» тесту

(тобто при $\varphi_{vi} = 1$) ми можемо забезпечити $D(k_v) = \sum_{i=1}^m \varphi_{vi} (1 - \varphi_{vi}) = 0$ і, отже, одержати $D(k) = D[M(k_v)]$ тобто позбутися «шуму», внесеного тестом.

Розглянуто кожен з категорій респондентів і виведені необхідні для неї формули.

Категорія « N »: респонденти S_v , $v=1, N_{S2}$ мають незадовільну оцінку хоча б по одній з навчальних програм Π_i ($i=1, m$). Їх матриця відповідей:

$$X_{vi} = \begin{cases} 1, & S_v \text{ відповів на питання по програмі } \Pi_i; \\ 0, & S_v \text{ не відповів на питання по програмі } \Pi_i. \end{cases}$$

Бал позитивних оцінок або кількість правильних відповідей респондента S_v дорівнює $k_v = \sum_{i=1}^m X_{vi}$; ($i=1, m = N_{\Pi}$); $k_v < N_{\Pi} = m$.

Категорія « U »: ті респонденти S_v , які не мають незадовільних оцінок з усіх предметів Π_i . Для категорії « U » відносні показники якості навчання дорівнюють

$$U_{R(U)} = 1 - \frac{\sum d_i}{m(N_S - N_{S2})}; \quad U_{Ri(U)} = 1 - \frac{d_i}{N_S - N_{S2}},$$

де d_i – кількість задовільних оцінок. Для категорії « U » величина $C = N_S - N_{S2} = N_{ir}$.

Категорія « K »: ті респонденти S_v , які мають з усіх Π_i гарні оцінки. Число респондентів у « K » дорівнює: $U_Z = N_S - N_{S2} - U_{4,3}$.

Категорія « D »: респондент S_v , який має з усіх Π_i оцінки «1» або «0». Ця категорія рівносильна категорії « N », якщо в ній немає оцінки, «добре» і «відмінно», тобто всі ті респонденти, які мають оцінки тільки «незадовільні» або «задовільні».

Таким чином, використовується двокритеріальний метод аналізу – по «абсолютній успішності» і по «якості успішності». Формально методи аналізу по кожному з критеріїв збігаються. Згідно з протоколом оцінок встановлюється чисельність респондентів вищої категорії для даного критерію. Якщо позначити це число через U_{CO} , то при аналізі абсолютної успішності це число дорівнює чисельності респондентів категорії « U », а при аналізі якості успішності – чисельності респондентів категорії « K ». Очевидно, що $U_{CO} = C - N_{CO}$. Для оцінки успішності в категорії « U »

$$C = N_S; N_{S2} = N_{CO}; N_{ir} = N_S - N_{CO} = C - N_{CO} = N_S - N_{S2} \rightarrow U_{CO}.$$

Для оцінки якості навчання в категорії « U »

$$N_{CO} = U_3; C = N_{ir}; U_{CO} \rightarrow N_{ir} - U_3 = C - N_{CO} = U_4;$$

де N_{CO} число респондентів нижчої категорії. Обчислюється відносне значення числа U_{CO} , тобто $U = \frac{U_{CO}}{C} = 1 - \frac{N_{CO}}{C}$, де C прирівнюється обліковому числу респондентів при аналізі абсолютної успішності або числу всіх «встигаючих» при аналізі якості успішності. Очевидно, що значення U і відповідні U_i використовуються для оцінки ефективності навчання (з. за даним критерієм) для безлічі $\{S_v\}$ у цілому і для окремих навчальних програм.

По якості навчання для категорії « U »

$$C = N_{ir} = N_S - N_{S2}; \quad U \rightarrow U_{R(U)} = 1 - \frac{\sum d_i}{mN_{ir}}; \quad U_i \rightarrow U_{Ri(U)} = 1 - \frac{d_i}{N_{ir}}.$$

Для категорії « $U + N$ » по загальній успішності:

$C = N_S$; \bar{d}_i – кількість незадовільних оцінок

$$d_i \rightarrow \bar{d}_i \left(d_{CO} \rightarrow \bar{d}_{CO} = \sum_{i=1}^m \bar{d}_i \right); \quad U \rightarrow U_K = 1 - \frac{\bar{d}_{CO}}{mN_S}; \quad U_i \rightarrow U_{iK} = 1 - \frac{\bar{d}}{N_S}.$$

Ефективність навчання даного контингенту оцінюється за кінцевим результатом, тобто за ступенем навченості респондента кожного S_v .

Разом з тим існуючі методи аналізу мають серйозний недолік – неможливість оцінки взаємодії основних елементів $(S, P \text{ і } W)$ системи навчання, тобто неможливість з'ясування хоча б причин відхилень ходу навчання від норми і вироблення, принаймні, напрямку керуючих впливів.

Далі розглянемо задачі синтезу математичної моделі взаємодії оператора і системи навчання.

Використання лінгвістичних правил «IF – THEN» дозволяє значно знизити обсяг експериментальних даних, необхідних для якісної ідентифікації. Для задач, де більш важливим є обґрунтування прийнятого рішення, мають перевагу нечіткі моделі типу Мамдани:

$$\bigcup_{p=1}^{k_j} \bigcap_{i=1}^n (x_i = a_{i,jp}) \rightarrow y = d_j, j = \overline{1, m} \quad (5)$$

де $a_{i,jp}$ – лінгвістичний терм, яким оцінюється перемінна x_i в рядку з номером jp ($p = \overline{1, k_j}$); k_j – кількість строк-кон'юнкцій, у яких вихід y оцінюється лінгвістичним термом d_j ; m – кількість термів, які використовують для лінгвістичної оцінки вихідної перемінної y .

Представлені правила віднесення оператора до деякого когнітивного типу. Для цього вводяться додаткові змінні P і L , що описують темп і рівень складності представлення інформації. Змінні L і P визначаються параметрами $N_1 \dots N_5$, що є результатами тестування продуктивних характеристик користувача $L = f_L(N_1, N_2, N_3)$; $P = f_P(N_4, N_5)$ з урахуванням правила віднесення оператора до деякого когнітивного типу. Завдання ідентифікації полягає в перебуванні нечіткої моделі F , що забезпечує мінімальне значення середньоквадратичної нев'язки:

$$R = \frac{1}{M} \sum_{j=1}^M (y_r - F(X_r))^2 \rightarrow \min$$

де $F(X_r)$ – значення виходу нечіткої моделі при значенні входів, заданих вектором X_r . Перебування структури і параметрів нечіткої моделі, що забезпечують мінімальне значення критерію, і є задачею ідентифікації.

Зіставити рівень знань («незнань») з позначеними в них об'єктами предметної області можна, оцінивши ступінь симетрії (асиметрії) цієї інформації і дійсності.

У підсумку ентропію знань пропонується оцінювати як

$$E = f(R, S_w)$$

де R – безліч зв'язків категорій бази знань; S_w – ентропія для предметної області W , визначеної як ступінь дезорганізації фантома дійсності, тобто його ентропію максимального значення x у вибірці. Партнери, що ведуть тематичний діалог, як і в інших випадках спілкування, переслідують деяку погоджену мету діалогу G_w для заданої предметної області W . Ідентифікатор чергового кроку діалогу (або питання відповідної ситуації)

визначимо як $i = f(G_w, \max(S_c), R_N)$, де G_w – мета діалогу; S_c – ентропія категорії; R_N – параметр, що визначає семантичну близькість категорій. Приведено варіаційно-ентропійний аналіз, який визначається умовою максимізації ентропії цієї системи S при фіксованих умовах, що відповідають її природі: $S = -\int f(x) \ln f(x) dx = \max \int E(x) f(x) dx = E$, $\int f(x) dx = 1$,

де S – ентропія; $E(x)$ – міра зусиль, необхідних для переведення даного елемента системи в стан x , E – міра зусиль, необхідних для переведення всієї системи в стан, що описується розподілом $f(x)$.

З ростом максимального вибіркового значення I випадкової величини x вона швидко збігається до кінцевої границі:

$$S = - \sum_{x_0}^I \frac{n(x)}{N} \log_2 \frac{n(x)}{N} = \frac{I}{\ln 2} \left[\ln \alpha^{-1} + \alpha^{-1} + I + \ln(x_0^{-\alpha} - I^{-\alpha}) + \frac{I + \alpha \ln(x_0^\alpha) / x_0^\alpha - \ln(I^\alpha) / I^\alpha}{\alpha \cdot \frac{x_0^{-\alpha} - I^{-\alpha}}{x_0^\alpha - I^\alpha}} \right]; \quad (6)$$

$$\lim_{\substack{I \rightarrow \infty \\ N \rightarrow \infty}} S = \frac{I}{\ln 2} (\ln \alpha^{-1} + \alpha^{-1} + I + \ln x_0),$$

де x_0 – мінімальне вибірконе значення x ; α – параметр розподілу Ципфа $n(x)$; N – обсяг вибірки.

Даний вираз є часткою ентропійної функції $H = \sum_i \mu_i \ln \mu_i$, де μ_i – міра на безлічі станів розглянутої системи для групи подій ($i = 1, 2, \dots$). Причому міра μ_i у загальному випадку не обов'язково повинна бути ймовірною, тобто задовольняти умові $\sum_i \mu_i = 1$.

Таким чином, ентропія S описується виразом

$$S = \sum_{x_0}^I \frac{n(x)}{N} \log_2 \frac{n(x)}{N}, \quad (7)$$

тобто без мінуса перед сумою в порівнянні з ентропією розподілу Ципфа.

При $N \rightarrow \infty$ і $x_0 = I$ маємо: $\lim_{\substack{I \rightarrow \infty \\ N \rightarrow \infty}} S = - \frac{I}{\ln 2} (\ln \alpha^{-1} + \alpha^{-1} + I)$.

Розкриємо $n(x)$ в загальному випадку одержимо:

$$S = - \frac{I}{\ln 2} \left[\ln \alpha^{-1} + \alpha^{-1} + I + \ln(x_0^{-\alpha} - I^{-\alpha}) + \frac{I + \alpha \cdot \frac{\ln(x_0^\alpha) / x_0^\alpha - \ln(I^\alpha) / I^\alpha}{x_0^{-\alpha} - I^{-\alpha}}}{\alpha} \right], \quad (8)$$

Так як $x_0^\alpha \square I$, а $I^{-\alpha} \square I$, то

$$S = - \frac{I}{\ln 2} \left[\ln \alpha^{-1} + \alpha^{-1} + I - \frac{I + \alpha}{\alpha} \times \frac{\ln(I^\alpha)}{I^\alpha} \right] \quad (9)$$

Вираз дозволяє оцінити ступінь дезорганізації фантома дійсності через його ентропію, тобто через одну змінну $x = I$. А це, у свою чергу, через максимальне значення x у вибірці, число елементів якої потенційно дорівнює N .

Розробка структурної схеми

Розглянемо методи параметричної ідентифікації математичних моделей оцінки знань операторів АСУ. Запропонуємо підхід, що припускає проведення з оператором серії активних експериментів, у продовж яких реєструються кінцеві результати процесу оцінювання.

З безлічі математичних моделей для опису ймовірності правильної відповіді в залежності від рівня знань θ_v v – того респондента та i труднощів i – того завдання обране однопараметричне нелінійне рівняння Георга Раша:

$$P_{vi} = \frac{e^{D(\theta_v - \beta_i)}}{1 + e^{D(\theta_v - \beta_i)}}, \quad (10)$$

де параметр D введений для того, щоб стандартизувати шкали ймовірностей для різних математичних моделей. Для v – того респондента функція Раша має вигляд:

$$P_v = \frac{e^{D(\theta_v - \beta)}}{1 + e^{D(\theta_v - \beta)}}. \quad (11)$$

Для одного i – го завдання функція Раша записується в такий спосіб:

$$P_i = \frac{e^{D(\theta - \beta_i)}}{1 + e^{D(\theta - \beta_i)}} \text{ при } \beta_i \rightarrow +\infty. \quad (12)$$

Загалом задача зводиться до визначення θ і β шляхом алгоритмічних дій і послідовних ітерацій, виходячи з експериментальних даних тестування:

$$X_v \rightarrow \theta_v, R_i \rightarrow \beta_i,$$

де X_v – сумарний бал v – того тестованого, а R_i – сума правильних відповідей для i – го завдання. Для виміру цих двох латентних змінних використовують логіт – модель або модель бінарного (3.дихотомічного) вибору U і ту саму одиницю виміру – логіт:

$$\ell_{\theta_v} = \ln \frac{P_v}{q_v}; P_v = \frac{X_v}{N}; q_v = 1 - p_v \text{ логіт рівня знань } v \text{ – го респондента,}$$

$$\ell_{\beta_i} = \ln \frac{q_i}{p_i}; p_i = \frac{R_i}{k}; q_i = 1 - p_i \text{ логіт рівня труднощів } i \text{ – і задачі,}$$

де ℓ_{θ_v} і ℓ_{β_i} – логіти рівня професійних знань v – того респондента і труднощі i – го завдання відповідно, P_v і q_v – частка правильних і неправильних відповідей для v – того респондента, p_i і q_i – частка правильних і неправильних відповідей по всіх респондентах для i – го завдання відповідно. Алгоритми обчислень розбиваються на ряд етапів: упорядковується матриця даних тестування; виробляється розрахунок початкових значень β_i^0 . Для всіх завдань по рівню обчислюються логіти труднощів завдань, потім їхнє

$$\bar{\ell}_{\beta} = \frac{\sum_i^k \ell_{\beta_i}}{k}.$$

середнє значення

У наступному етапі переносять центр розподілу логітів труднощів на завдання: $\beta_i^0 = \ell_{\beta_i} - \bar{\ell}_{\beta}$.

При формуванні тесту в нього потрібно включати завдання з різними значеннями θ_v і β_i . Якщо тестована група однорідна за рівнем професійних знань, то в тест необхідно брати завдання з великою крутістю характеристики. Якщо ж група неоднорідна, то в тест включаються завдання з малою крутістю, але при цьому треба намагатися, щоб криві завдань не перетиналися і були розташовані по всьому простору. Величина α_i обчислюється за допомогою бісеріального коефіцієнта кореляції між балами i – го завдання і сумою індивідуальних балів усіх респондентів:

$$\alpha_i = \frac{(r_{bis})_i}{\sqrt{1 - (r_{bis})_i^2}}.$$

Для системи тестового контролю професійних знань бісеріальний коефіцієнт кореляції описується залежністю:

$$(r_{bis})_i = \frac{(\bar{X}_1)_i - (\bar{X}_0)_i}{S_x} \cdot \frac{(n_1)_i (n_0)_i}{N_i \cdot N \sqrt{N(N-1)}}, \quad (13)$$

де $(\bar{x}_1)_i$ – середній сумарний бал тих N_i респондентів, що на i -те завдання дали правильну відповідь ($X_{vi}=1$), $(\bar{x}_0)_i$ – теж саме для тих респондентів, що на i -те завдання дали неправильну відповідь ($X_{vi}=0$); n_{1i} і n_{0i} – кількість правильних і неправильних відповідей на i -те завдання; u_i – ордината функції нормального розподілу в точці; S_x – вибіркова дисперсія.

Розробка математичної моделі рівняння зв'язку для системи навчання починається з після оптимізаційного етапу. Після оцінки погодженості і конформності питання виявилися достатніми для віднесення тесту до класу однорідних, з'являється можливість використання такого оптимального тесту в якості системного при програмованому контролі знань. Основою для такої можливості є визначені функціональні зв'язки параметрів тесту (тобто деяких констант, що задають модель тесту) з тестовими показниками (тобто величинами, що спостерігаються в процесі тестування: числами вірних відповідей на питання тесту). Зрозуміло, що ці зв'язки є детермінованими тільки стосовно деякої генеральної сукупності респондентів, чисельність якої N достатня для застосування граничних теорем, тобто оперування поняттями ймовірностей, математичних сподівань і т.п. замість частот, середніх значень і т.д. Як модель тесту використовується цілком визначений функціональний зв'язок ймовірностей для випадкової події $X_{vi}=1$ (v -й респондент дав вірну відповідь на i -те питання тесту):

$$\begin{cases} P[X_{vi}=1/A_v] = S; & P[X_{vi}=1/\bar{A}_v] = P_i, \\ P[X_{vi}=1] = Z_v * S + (1 - Z_v)P_i = \hat{O}_{vi}, \end{cases} \quad (14)$$

де параметри \hat{A} -тесту S і P_i є константами і визначають S – ймовірність вірної відповіді за умови знання; P_i – ймовірність вірної відповіді за умови незнання, тобто ймовірність угадування. Обидва ці показники є випадковими величинами і мають асимптотично нормальні розподіли. Розглянемо спочатку розподіл ваги питання B_i з урахуванням однорідності тесту, для будь-якого B_i маємо

$$\begin{cases} \hat{M}(B_i) = (S - P_i) \sum_{v=1}^N Z_v + NP_i; \\ \hat{D}(B_i) = (S - P_i)(1 - 2P_i) \sum_{v=1}^N Z_v - (S - P_i)^2 \cdot \sum_{v=1}^N Z_v^2 + NP_i(1 - P_i). \end{cases} \quad (15)$$

Вхідні у виразі (15) для $\hat{M}(B_i)$ і $\hat{D}(B_i)$ суми $\sum_v Z_v$ і $\sum_v Z_v^2$ розглядаються як параметри розподілу рівнів знань у генеральній сукупності N . Дійсно, кожний з респондентів у цій сукупності володіє деяким Z_v (тобто визначеною ймовірністю знання). Звернемося далі до другого тестового показника, так званому бала тесту K . Кожна з умовних ймовірностей $k_v = j$ у зв'язку з однорідністю тесту відшукується у вигляді:

$$\begin{cases} P[k_v=j/A_v] = C_m^j S^j (1 - S)^{m-j}; \\ P[k_v=j/\bar{A}_v] = C_m^j P^j (1 - P_i)^{m-j}, \end{cases} \quad (16)$$

що дозволяє знайти ймовірність $P[k=j]$ для бала тесту у вигляді: $P[k=j] = C_m^j S^j (1-S)^{m-j} M(Z) + C_m^j P_i^j (1-P_i)^{m-j} [1-M(Z)]$. Отже, при відомих параметрах розподілів ваги питання $M(B_i)$ і $D(B_i)$, а також бала тесту $M(K)$ і $D(K)$ є можливість скласти рівняння зв'язку цих величин з параметрами тесту S і P_i , а також параметрами розподілу рівнів знань $M(Z)$ і $D(Z)$ у вигляді системи рівнянь:

$$\begin{cases} M(K) = M(S - P_i)M(Z) + MP_i; \\ M(B_i) = N(S - P_i)M(Z) + NP_i; \\ D(K) = [1 + (M - 1)(S + P_i) - 2MP_i]M(S - P_i)M(Z) - \\ - M^2(S - P_i)^2 M(Z) + MP_i(1 - P_i); \\ D(B_i) = M(Z)(S - P_i)(1 - 2P_i) + NP_i(1 - P_i) - \\ - N(S - P_i)^2 D(Z) - N(S - P_i)^2 M^2(Z). \end{cases} \quad (17)$$

Перші два рівняння в системі не є незалежними, тому що з теорії дихотомічних тестів випливає очевидне співвідношення:

$$\frac{M(K)}{M} = \frac{M(B_i)}{N} = \frac{M(B)}{MN} = \alpha, \quad (18)$$

$$\text{де } B_i = jN_j = \sum_{v=1}^N X_{vi}; \quad B = \sum_{i=1}^m B_i$$

– вага тесту. Таким чином, система дозволяє розв'язок по відношенню тільки трьох невідомих з $S; P_i; M(Z); D(Z)$, і то за умови, що попередньо знайдено математичні сподівання і дисперсії тестових показників K і B_i . Крім того, необхідно врахувати, що питання оцінок довірчих інтервалів вирішені у теоретичній статистиці лише для моментів двох перших порядків, хоч оцінки максимальної правдоподібності можна знайти для моментів будь-якого порядку.

У цих умовах переважнішим виявляється шлях використання цієї системи рівнянь, що дозволяє при відомому P_i відшукати параметр тесту S у вигляді

$$S = \frac{D(K) + M^2(K)(1 - P_i + MP_i)}{(M - 1)[M(K) - MP_i]}$$

Що ж стосується ймовірності угадування P_i , то цей параметр діагностичного тесту можна вважати рівним $P_i = \frac{1}{H}$, де $H = 2, 3, \dots, n$ – число відповідей до питань тесту. За умови підвищених вимог до вірогідності оцінки S можна знаходити параметр P_i методом рандомізації підсумків тесту. Неминучість використання, що відзначалася $M(K)$ вище, $D(K)$ замість їхніх оцінок приводить до необхідності розгляду $D(K) = f[M(K)]$ для дослідження $S = const$ функції $P_i = const$.

Для границь довірчих інтервалів дисперсії D'_K і $D''_K > D'_K$; D'_B і $D''_B > D'_B$ і відповідних коефіцієнтів довіри $Q_K = 1 - \beta_K$, $Q_B = 1 - \beta_B$ можна скласти рівняння зв'язку у вигляді:

$$\left\{ \begin{array}{l} P_{m-1}\{X_B''\} = 1 - \frac{\beta_B}{2}; \quad P_{m-1}\{X_B'\} = \frac{\beta_B}{2}; \\ X_B'' = \frac{S_B^2}{D_B''}(M-1); \quad X_B' = \frac{S_B^2}{D_B'}(M-1); \\ \hat{O}_0 \left\{ \frac{X_K'' - N + 1}{\sqrt{2N-2}} \right\} = \frac{\beta_K}{2}; \quad \hat{O}_0 \left\{ \frac{X_K' - N + 1}{\sqrt{2N-2}} \right\} = 1 - \frac{\beta_K}{2}; \\ X_K'' = \frac{S_K^2}{D_K''}(N-1); \quad X_K' = \frac{S_K^2}{D_K'}(N-1). \end{array} \right. \quad (19)$$

Структурна схема розробленої системи зображена на рисунку 1. На ній показано структуру системи ідентифікації професійних знань операторів автоматизованих систем управління.



Рисунок 1 – Структурна схема системи

Висновки. У статті наведені теоретичне узагальнення й рішення наукового завдання дослідження методів ідентифікації професійних знань операторів автоматизованих систем управління. Рішення даного завдання полягало у вирішенні наступних задач: Був проведений огляд існуючих систем ідентифікації професійних знань операторів автоматизованих систем управління; Досліджена система ідентифікації професійних знань операторів автоматизованих систем управління; На основі отриманих результатів досліджень створена програмна реалізація системи ідентифікації професійних знань операторів автоматизованих систем управління; Розроблені під час виконання випускної кваліфікаційної роботи за другим (магістерським) рівнем вищої освіти алгоритми дозволяють успішно вирішувати завдання ідентифікації професійних знань операторів автоматизованих систем управління. Проведено аналіз предметної галузі в ході якого були виявлені об'єкти, взаємодія яких носить істотний характер для функціональної діяльності

предметної галузі, і їхні основні характеристики; побудована алгоритм і вибраний середовище розробки.

Список літератури

1. Smirnov, O., Odarchenko, R., Smirnova, T., Bondar, S., Volosheniuk, D. «Optimal Structure Construction of Private 5G Network for the Needs of Enterprises». *Lecture Notes on Data Engineering and Communications Technologies*, 2023, 178, pp. 208–223.
2. Smirnov, O., Karapetyan, A., Fedorov, E., «Creating Neural Network and Single Solution Human-Based Metaheuristic Methods of Solving the Traveling Salesman Problem». *CEUR Workshop Proceedings, Volume 3312*, 2022, pp. 47-58.
3. Smirnov O., Kuznetsov A., Kryvinska N., Kiian A., Kuznetsova K. «Full Non-Binary Constant-Weight Codes». *SN Computer Science*, Vol 2, 337, 2021. <https://doi.org/10.1007/s42979-021-00739-w>.
4. Smirnov O., Kovalenko O., Kovalenko A., Kavun S. «Quantitative Risk Assessment Method Development in the Context of the SDLC-model». *2021 IEEE 8th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T)*, 2021, pp. 203-208, doi: 10.1109/PICST54195.2021.9772143
5. Smirnova T., Gnatyuk S., Berdibayev R., Avkurova Zh., Iavich M. «Cloud-Based Cyber Incidents Response System and Software Tools». *Communications in Computer and Information Science*, 2021, vol 1486. Springer, Cham. pp 169-184.
6. Smirnov, O., Kuznetsov, A., Potii, O., Poluyanenko, N., Stelnyk, I., Mialkovsky, D. «Combining and filtering functions in the framework of nonlinear-feedback shift register». *International Journal of Computing*; 2020, Volume 19, Issue 2 – Research Institute for Intelligent Computer Systems – 2020. – P. 247-256.
7. Smirnov O., Kuznetsov A., Kiian A., Kuznetsova T. «Non-binary constant weight coding technique». *CEUR Workshop Proceedings. Volume 2740*, 2020, Pages 102-114.
8. Smirnov O., Kuznetsov A., Kiian A., Cherep A., Kanabekova M., Chepurko I. «Testing of code-based pseudorandom number generators for post-quantum application». *2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, Ukraine, Kyiv, May 14-18. 2020. P. 172-177.
9. Smirnov, O., Shekhanin, K., Kuznetsov, A., Krasnobayev, V. «Detecting Hidden Information in FAT». *International Journal of Computer Network and Information Security (IJCNIS)*. Vol. 12, No. 3, 2020. PP.33-43.
10. Smirnov, O., Drieieva, H., Drieiev, O., Simakhin, V., Bondar, S., Odarchenko, R. «Managing multifractal properties of the binary sequence generated with the Markov chains», *CEUR Workshop Proceedings Volume 2608*, 2020, Pages 633-645.
11. Smirnov O. Kuznetsov A., Zaichenko Yu., Pastukhov M., Oleshko O., Kuznetsova K., «Formation of Discrete Signals with Special Correlation Properties». *International Conference on Information and Telecommunication Technologies and Radio Electronics, UkrMiCo 2019*; Odessa; Ukraine; 9-13 September 2019. P.22-28.
12. Smirnov, O., Kuznetsov, A., Kolovanova, I., Kuznetsova, T., «Noise immunity of the algebraic geometric codes». *International Journal of Computing*; 2019, Volume 18, Issue 4 – Research Institute for Intelligent Computer Systems – 2019. – P. 393-407.
13. Smirnov, O., Kuznetsov, A., Reshetniak, O., Ivko, N., Katkova, T., Kuznetsova, T., «Generators of Pseudorandom Sequence with Multilevel Function of Correlation». *2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T)*, Kyiv, Ukraine, 8 – 11 October 2019 . P.517-522.
14. Smirnov, O., Krasnobayev, V., Yanko, A., Kuznetsova, T. «Methods of nulling numbers in the system of residual classes». *CEUR Workshop Proceedings, Vol 2588*, P. 90-106, 2019.
15. Kuznetsova, T., «Code-Based Schemes for Post-Quantum Digital Signatures», *10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2019*; Metz; France; 18-21 September 2019. P. 707-712.
16. Smirnov, O., Kuznetsov, A., Stefanovych, O., Gorbenko, Y., Krasnobayev, V., Kuznetsova K. «Information Hiding Using 3D-Printing Technology», *10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2019*; Metz; France; 18-21 September 2019. P.701-706.
17. Smirnov, O., Kuznetsov, A., Kovalchuk, D., Averchev, A., Pastukhov, M., Kuznetsova, K., «Formation of Pseudorandom Sequences with Special Correlation Properties», *2019 3rd International Conference on Advanced Information and Communications Technologies, AICT -2019/ Lviv, Ukraine, 2-6 July, 2019*, P. 395-399.
18. Вінтенко Б.Ю., Смірнов О.А., Коваленко А.С., Смірнов С.А., Буравченко К.О. «Дослідження вимог міжнародних стандартів ІЕС60880 та ІЕС62138 з розробки програмного забезпечення інформаційно-керуючих систем АЕС, важливих для безпеки». *Системи управління, навігації та зв'язку*, 2023, вип. 3(73), С. 155-166.
19. Вінтенко, Б., Миронець, І., Смірнов, О., Кравчук, О., Козірова, Н., Савеленко, Г., Коваленко, А. «Дослідження вимог та аналіз кібербезпеки програмного забезпечення інформаційно-керуючих систем АЕС, важливих для безпеки». *Кібербезпека: освіта, наука, техніка*. 2024. №3(23), С. 111-131.

20. Вінтенко Б.Ю., Смірнов О.А., Коваленко О.В., Смірнов С.А., Коваленко А.С. «Дослідження нормативних документів та галузевих стандартів розробки програмного забезпечення комп'ютерних систем управління АЕС, важливих для безпеки». Системи управління, навігації та зв'язку, 2023, вип. 2(72), С. 170-178.
21. Аль-Мудхафар Акіл Абдулхуссейн М., Смірнова Т.В., Буравченко К.О., Смірнов О.А. «Метод оцінки та підвищення користувальницького досвіду абонентів в програмно-конфігурованих мережах на основі використання машинного навчання». Сучасні інформаційні системи, 2023, том 7, № 2, С. 49-56.
22. Вінтенко Б.Ю., Смірнов О.А., Коваленко О.В., Смірнов С.А. «Дослідження нормативної документації та стандартів розробки програмного забезпечення комп'ютерних систем управління АЕС, важливих для безпеки». VI міжнародна науково-практична конференція «Інформаційна безпека та комп'ютерні технології», м. Кропивницький. 20-21 квітня 2023 р. – Кропивницький: ЦНТУ. – 2023. – С. 35-36.
23. Смірнов, О.А., Усік П.С., Полігенько О.О., Одарченко Р.С., Терещенко Л.Ю. «Інформаційна технологія та програмне забезпечення для підвищення ефективності планування підсистеми базових станцій стільникового зв'язку». Проблеми телекомунікацій. № 1(26). С. 83-96. 2020.