

Методологія створення інтелектуального інтерфейсу для систем прийняття рішень

Євген Є. Федоров

Кафедра спеціалізованих комп'ютерних систем
Донецька академія автомобільного транспорту, Донецьк
fee75@mail.ru

Анотація

У тезах запропонована методологія створення інтелектуального інтерфейсу для систем прийняття рішень, що використовує вузькоспеціалізований технічний словник ОПР, що управляє технологічним процесом.

1. Вступ

Пропонована методологія присвячена принципам розробки службової підсистеми інтелектуального спілкування людини з комп'ютером, що управляє технологічним процесом. Службова підсистема може застосовуватися в автоматичних й автоматизованих режимах. При керуванні конкретними об'єктами використовується технічний словник, що заснований на вузькоспеціалізованій технології керування процесами, які протікають на об'єкті. Основою службової підсистеми є інтелектуальний інтерфейс, що базується на пропонованій методології вузькоспеціалізованого словника. Функціональне призначення системи полягає у двох режимах – навчання й керування.

Функціонально службова підсистема в режимі навчання припускає виконання декількох операцій. Тому що в керуванні агрегатом бере участь кілька операторів, то на першому етапі створюються еталони фонем, що відображають мовні особливості кожного оператора (ОПР). При прийомі зміни для санкціонування допуску до керування об'єктом оператор вимовляє ключові слова пароля. Відповідно до цього в базу даних завантажуються еталони, які відповідають фонемам даного ОПР. Це означає початок готовності системи до роботи.

Функціонально службова підсистема в режимі керування припускає виконання декількох операцій. ОПР, що одержав допуск до керування об'єктом, протягом зміни через інтелектуальний інтерфейс видає повідомлення про стан об'єкта або керуючі команди. У режимі автоматичного керування значення частина повідомлення або команди розпізнається шляхом зіставлення з еталонами й передається в основну частину системи керування процесом, що приймає відомості про стан об'єкта й виробляє керуючий вплив. В автоматизованому режимі на підставі запропонованої методології формується рішення, що передається ОПР для виконання. Після виконання команди по зворотному зв'язку через інтелектуальний інтерфейс ОПР одержує мовне повідомлення про виконання команди.

Сутність методології створення інтелектуального інтерфейсу й службової підсистеми містить у собі рішення наступних питань:

1. У режимі навчання:

– для фіксації мовних особливостей ОПР і формування еталонів фонем у службовій підсистемі формується система векторів ознак і наповнення їхніми кількісними оцінками;

– після визначення чисельних значень компонент векторів ознак у базу даних заносяться еталони фонем;

– на підставі створених еталонів фонем навчається різноманіття нейромереж, що представляє модель службової підсистеми. На підставі еталонів формуються структури мереж. Для обраних структур нейромереж формується різноманіття нейромереж, що відповідає різноманіттю мовних особливостей ОПР;

– за допомогою генетичного алгоритму здійснюється ідентифікація «у великому» й «у малому» різноманіття нейромереж.

У результаті ідентифікації різноманіття нейромереж службова підсистема готова до роботи.

2. У режимі керування:

– при керуванні процесом від ОПР через інтелектуальний інтерфейс у службову підсистему надходить повідомлення про стан об'єкта або команди на керування. Відповідно до запропонованої методології визначаються кількісні характеристики векторів ознак поданих повідомлень або команд, які розпізнаються за допомогою різноманіття нейромереж шляхом зіставлення з еталонами;

– розпізнаний зміст повідомлення через інтелектуальний інтерфейс надходить в основну частину системи керування для оцінки стану об'єкта;

– розпізнана команда в режимі автоматичного керування передається в основну частину системи керування процесом, що виконує її, а в автоматизованому режимі на підставі запропонованої методології формується рішення, що передається ОПР для виконання.

Після описаних в узагальненому виді режимів навчання й керування далі розглядаються складові частини запропонованої методології.

2. Формування системи векторів ознак

У режимі навчання для фіксації мовних особливостей ОПР і формування еталонів фонем у службовій підсистемі формується система з 14 векторів ознак:

- коефіцієнти КЛП (кодування з лінійним прогнозуванням);
- коефіцієнти відбиття КЛП;
- автокореляція КЛП;
- кепстр КЛП;
- площі поперечних перерізів акустичної труби;
- нормована автокореляція;
- нормований енергетичний спектр КЛП;
- міри контрастності КЛП;

- нормований енергетичний спектр ДПФ (дискретного перетворення Фур'є);
- міри контрастності ДПФ;
- міри контрастності ДВП (дискретного вейвлет-перетворення);
- міри контрастності БВП (безперервного вейвлет-перетворення);
- кількість імпульсів рівної довжини;
- MFCC (мел-частотні кепстральні коефіцієнти).

Сформовані вектора ознак відбивають частотні, автокореляційні й енергетичні характеристики мови ОПР, його мовного тракту й слухового апарата й використовуються для створення еталонів.

3. Формування еталонів

Для виділених у попередньому розділі компонентів векторів ознак визначаються кількісні значення цих компонентів, на підставі чого в службовій підсистемі формуються множини еталонів фонем виду

$$E_u = \{E_{ui}\}, E_{ui} = (\mathbf{m}_{ui}, \mathbf{C}_{ui}), u \in \overline{1,14}, i \in \overline{1, N_u^{(1)}}, \quad (1)$$

$$\mathbf{m}_{ui} = (m_{ui1}, \dots, m_{uik}, \dots, m_{uiN_u}), \quad (2)$$

$$\mathbf{C}_{ui} = \text{diag}(\sigma_{ui1}^2, \dots, \sigma_{uik}^2, \dots, \sigma_{uiN_u}^2), \quad (3)$$

де \mathbf{m}_{ui} – вектор математичних очікувань чисельних значень компонент u -го вектора ознак розмірності N_u ,

\mathbf{C}_{ui} – діагональна матриця дисперсій чисельних значень компонент u -го вектора ознак розмірності $N_u \times N_u$,

N_u – довжина u -го вектора ознак,

$N_u^{(1)}$ – кількість еталонів для u -го вектора ознак.

На підставі сформованих еталонів фонем створюються структури нейромереж.

4. Формування структури нейромереж як моделі інтелектуального інтерфейсу

На підставі сформованих еталонів як модель інтелектуального інтерфейсу обране різноманіття нейромереж, і для цих мереж визначені структури на підставі ймовірнісної мережі із двома шарами.

Нейромережі формуються по наступному принципі. Компонентам вектора мовного сигналу, які розглянуті в другому розділі, відповідають нейрони вхідного шару; еталонам певної множини (1), які сформовані в третьому розділі, відповідають нейрони схованого шару; номеру ОПР (при фонемній ідентифікації ОПР) або номеру фонемі (при фонемному розпізнаванні повідомлень або команд) відповідають нейрони вихідного шару.

Моделі для різноманіття нейромереж представлені у вигляді

$$\bar{f}_u(\mathbf{x}) = \{f_{ij}(\mathbf{x})\}, f_{ij}(\mathbf{x}) = \frac{1}{n_{ij}} \sum_{i=1}^{N_u^{(1)}} w_{ij} G_{ui}(\mathbf{x}), j \in \overline{1, N_u^{(2)}}, \quad (4)$$

де $G_{ui}(\mathbf{x})$ є багатомірним розподілом Гауса

$$G_{ui}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^{N_u} \det \mathbf{C}_{ui}}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{m}_{ui})^T \mathbf{C}_{ui}^{-1}(\mathbf{x} - \mathbf{m}_{ui})\right), \quad (5)$$

де \mathbf{x} – вхідний вектор ознак розмірності N_u ,

N_u – кількість нейронів вхідного шару,

$N_u^{(1)}$ – кількість нейронів у першому шарі,

$N_u^{(2)}$ – кількість нейронів у другому шарі,

$w_{ij} \in \{1,0\}$ – вагові коефіцієнти.

Для створеного різноманіття нейромереж необхідно провести ідентифікацію «у великому» й «у малому».

5. Ідентифікація «у великому» й «у малому» різноманіття нейромереж

У зв'язку з тим, що мовні особливості оператора можуть відрізнятися від його фізичного стану, службова підсистема передбачає наявність набору еталонів одного оператора на ті самі команди й повідомлення.

Для ідентифікації «у великому» й «у малому» різноманіття мереж використовується генетичний алгоритм.

Під ідентифікацією «у великому» розуміється визначення оптимальної структури (мінімальної кількості нейронів у першому шарі $N_u^{(1)}$), що забезпечує мінімальний час прогнозу по моделі (4).

Під ідентифікацією «у малому» розуміється визначення чисельних значень компонент векторів (2) і матриць (3), які забезпечують максимальну ймовірність прогнозу й мінімальну середньоквадратичну помилку прогнозу.

Генетичний алгоритм для кожної нейромережі містить у собі наступні блоки:

- подання множини еталонів нейромережі (1) у вигляді особини й створення з різних варіантів множини (1) вихідної популяції;
- функціонал мети (фітнес-функція);
- відбір кращих варіантів множини еталонів (оператор репродукції);
- комбінування двох варіантів множини еталонів (оператор кросінгвера);
- модифікація еталона у варіантах множини еталонів (оператор мутації);
- вибір з отриманих варіантів множини еталонів (оператор редукції);
- умова зупинки.

5.1. Подання особин і створення вихідної популяції

У першому блоці генетичного алгоритму множин еталонів нейромережі (1) представляються у вигляді s -ї особини з речовинними генами h_{is} й з різних варіантів множини (1) створюється вихідна популяція $H_u = \{h_{is}\}$,

$$h_{is} = ((l_{u11} + s * \Delta m_{u11}, \dots, l_{u1N_u} + s * \Delta m_{u1N_u}), \dots, \quad (6)$$

$$(l_{uN_u^{(1)}1} + s * \Delta m_{uN_u^{(1)}1}, \dots, l_{uN_u^{(1)}N_u} + s * \Delta m_{uN_u^{(1)}N_u}),$$

$$(l_{u11} + s * \Delta \sigma_{u11}, \dots, l_{u1N_u} + s * \Delta \sigma_{u1N_u}), \dots,$$

$$(l_{uN_u^{(1)}1} + s * \Delta \sigma_{uN_u^{(1)}1}, \dots, l_{uN_u^{(1)}N_u} + s * \Delta \sigma_{uN_u^{(1)}N_u}),$$

$s \in \overline{1, |H_u|}$,

$$\Delta m_{uik} = \frac{rx_{uik} - l_{uik}}{|H_u|}, \Delta \sigma_{uik} = \frac{rx_{uik} - l_{uik}}{|H_u|}, k \in \overline{1, N_u},$$

де $|H_u|$ – потужність популяції, lx_{ik} , rx_{ik} – ліва й права границі значень k -го ознаки i -го еталона u -ї нейромережі.

Сформовані вектора (6) використовуються у функціоналі мети (фітнес-функції).

5.2. Фітнес-функція

Як функціонал мети (фітнес-функції) методологія передбачає використання трьох критеріїв:

1. Критерій надійності, що для даного випадку означає вибір такої множини еталонів виду (6) для u -ї нейромережі, які доставляють максимум імовірності розпізнавання (відносини кількості правильно розпізнаних ОПР, повідомлень або команд по моделі (нейромережі) до їхньої загальної кількості)

$$F(h_{us}) = \frac{1}{P} \sum_{p=1}^P I(\operatorname{argmax}_j f_{up}(h_{us}), \operatorname{argmax}_j d_{jp}) \rightarrow \max_{E_u}, \quad (7)$$

$$\min h_{us} < h_{us} < \max h_{us},$$

$$I(a,b) = \begin{cases} 1, & a=b \\ 0, & a \neq b \end{cases}$$

де $\max h_{us}$, $\min h_{us}$ – максимальне й мінімальне значення i -го еталона,

$f_{up}(h_{us})$ – виходи нейромережі,

d_{jp} – виходи об'єкта,

P – кількість тестових реалізацій.

2. Критерій точності, що для даного випадку означає вибір такої множини еталонів виду (6) для u -ї нейромережі, які доставляють мінімум середньоквадратичної помилки прогнозу (квадрата різниці виходу нейромережі й виходу об'єкта)

$$F(h_{us}) = \frac{1}{P} \sum_{p=1}^P \sum_{j=1}^{M^{(2)}} (f_{up}(h_{us}) - d_{jp})^2 \rightarrow \min_{E_u}, \quad (8)$$

$$\min h_{us} < h_{us} < \max h_{us}.$$

3. Критерій швидкодії, що для даного випадку означає вибір оптимальної структури u -ї нейромережі (мінімальної кількості нейронів у першому шарі $N_u^{(1)}$), що забезпечує мінімальний час прогнозу по моделі (4)

$$F(h_{us}) = T \rightarrow \min_{N_u^{(1)}}, \quad (9)$$

$$\min N_u^{(1)} < N_u^{(1)} < \max N_u^{(1)}.$$

де $\max N_u^{(1)}$, $\min N_u^{(1)}$ – максимальна й мінімальна кількість нейронів у першому шарі.

Функціонал мети (7) або (8) використовується для відбору кращих варіантів множині еталонів виду (6).

5.3. Оператор репродукції

Для відбору кращих варіантів множині еталонів виду (6) як оператор репродукції використовується пропорційний відбір (рулетка).

Число копій для кожного s -го варіанта множині еталонів виду (6) залежить від величини його фітнес-функції й визначається у вигляді

$$V_{us} = \operatorname{round} \left(\frac{F(h_{us})}{\sum_{q=1}^{|H_u|} F(h_{uq})} |H_u| \right), \quad s \in \overline{1, |H_u|} \quad (10)$$

Щоб пошук оптимальної множини еталонів виду (6) був спрямованим, виробляється комбінування відібраних варіантів.

5.4. Оператор кросингвера

Для комбінування двох варіантів множині еталонів виду (6), відібраних оператором репродукції, як оператор кросингвера використовується дискретна рекомбінація.

Обидва варіанти множині еталонів вибираються випадково. Потім для кожного еталона множині з рівною ймовірністю вибирається перший або другий батько. Здійснюється схрещування, і виробляються два нащадки. Імовірність кросингвера - 0.5.

Для глобального пошуку оптимальної множини еталонів виду (6) необхідно підвищити розмаїтість варіантів.

5.5. Оператор мутації

Для забезпечення розмаїтості варіантів множині виду (6) після кросингвера використовується оператор мутації.

Випадково вибирається варіант множині виду (6). Потім випадково вибирається компонент вектора (6), до якої додається крок мутації Δ .

Для обчислення кроку мутації використовується алгоритм імітації відпала

$$\Delta = \begin{cases} (\max h_{usz} - h_{usz}) \left(1 - r^{\left(\frac{1-t}{T} \right)^2} \right), & r \geq 0 \\ (h_{usz} - \min h_{usz}) \left(1 - (-r)^{\left(\frac{1-t}{T} \right)^2} \right), & r < 0 \end{cases}, \quad (11)$$

де $\max h_{usz}$, $\min h_{usz}$ – максимальне й мінімальне значення z -го компонента,

t – номер ітерації,

T – максимальна кількість ітерацій,

r – випадкове число, $r \in [-1, 1]$.

Для цього алгоритму ймовірність мутації зменшуватися з кількістю ітерацій

$$P_m = P_0 \exp(-1/t), \quad (12)$$

де $P_0 = 0.05 \div 0.1$.

Далі виробляється вибір кращих варіантів з отриманих.

5.6. Оператор редукції

Для вибору з отриманих варіантів множині еталонів виду (6) (об'єднання вихідної популяції з результатами кросингвера й мутації) як оператор редукції використовується селекційна схема.

Старі й нові варіанти поєднуються в одна множин (популяцію) і впорядковуються по убаванню значення цільової функції. Відбираються перших $|H_u|$ варіантів (особин).

5.7. Умова зупинки

Пропонується наступна умова завершення генетичного алгоритму

$$1 - \max_s F(h_{is}) < \varepsilon \vee t \geq T, \quad (13)$$

Значення T обчислюється експериментально, ε задаються.

Якщо $t \geq T$, але $1 - \max_s F(h_{is}) > \varepsilon$, те відбувається додавання еталонів у множин еталонів виду (6) і генетичний алгоритм знову запускається.

Результатом роботи генетичного алгоритму є множина еталонів (особина) h_{is^*} , що задовольняє функціоналам мети (7)-(9).

6. Чисельне дослідження

Для одержання доступу до керування системою в рамках запропонованої методології було проведено чисельне дослідження розробленої службової підсистеми для вибору кращих векторів ознак.

Результати чисельного дослідження ідентифікації ОПР на основі імовірнісної нейромережі по фонемі |a|, що виділена із ключового слова, наведені в табл.1. Як треба з табл.1, кращими векторами ознак є міри контрастності БВП (на основі вейвлету Морле) і коефіцієнти відбиття КЛП. Обое ці вектора різної фізичної природи – перша враховує особливості слухового апарата, друга – мовної апарат, тому пропонується їхнє спільне використання.

Аналогічні результати були отримані при розпізнаванні повідомлень і команд.

Таблиця 1: Ідентифікація ОПР по фонемі |a| ключового слова на основі імовірнісної нейромережі

Вектора ознак	Імовірність розпізнавання
коефіцієнти КЛП	0.84
коефіцієнти відбиття КЛП	0.96
автокореляція КЛП	0.76
кепстр КЛП	0.8
площі поперечних перерізів акустичної труби КЛП	0.56
нормована автокореляція	0.44
нормований енергетичний спектр КЛП	0.1
міри контрастності КЛП	0.22
нормований енергетичний спектр ДПФ	0.72
міри контрастності ДПФ	0.68
міри контрастності ДВП	0.66
міри контрастності БВП	0.98
нормована кількість імпульсів рівної довжини MFCC	0.2
	0.66

7. Висновки

У тезах запропонована методологія створення інтелектуального інтерфейсу для систем прийняття рішень, що використовує вузькоспеціалізований технічний словник ОПР, що веде технологічний процес.

Службова підсистема може застосовуватися в автоматичних й автоматизованих режимах і передбачає режим навчання й режим керування.

Для режиму навчання здійснюється формування системи векторів ознак мови ОПР; створення на їхній основі еталонів фонем; визначення структур і моделей різноманіття нейромереж, які містять ці еталони; ідентифікація «у великому» й «у малому» різноманіття нейромереж за допомогою генетичного алгоритму.

Режим керування передбачає зіставлення векторів ознак поданих повідомлень або команд із еталонами різноманіття мереж і визначення, яке повідомлення або команда відповідає вектору ознак; визначення правочинності розпізнаного повідомлення або команди відповідно до функціонала мети; вироблення рішення за результатами розпізнавання.

Для одержання доступу до керування системою в рамках запропонованої методології було проведено чисельне дослідження розробленої службової підсистеми для вибору кращих векторів ознак. За результатами чисельного дослідження минулого обрані два вектори ознак – коефіцієнти відбиття КЛП і міри контрастності БВП, які дають найбільшу ймовірність правильної ідентифікації ОПР по фонемах і мають різну фізичну природу, що дозволяє зробити систему ознак неоднорідної. Аналогічні результати були отримані при розпізнаванні повідомлень і команд.

Службова підсистема, що розроблена на основі запропонованої методології, може використатися для інтелектуального керування агрегатами в різних галузях (вугільної й металургійної промисловості, авіабудування й суднобудування, нафтопроводів і газопроводів, атомних, теплових і гідроелектростанцій й ін.).

8. Література

- [1] Винцюк, Т.К., *Анализ, распознавание и интерпретация речевых сигналов*, Наук. думка, Киев, 1987, 261 с.
- [2] Rabiner, L.R., Jang, B.H., *Fundamentals of speech recognition*, Prentice Hall PTR, Englewood Cliffs, New Jersey, 1993, 507 p.
- [3] Хайкин, С., *Нейронные сети: полный курс*, Издательский дом «Вильямс», Москва, 2006, 1104 с.
- [4] Рутковская, Д., Пилиньский, М., Рутковский, Л., *Нейронные сети, генетические алгоритмы и нечеткие системы*, Горячая линия – Телеком, Москва, 2006, 452 с.
- [5] Федоров, Е.Е., *Методика формирования акустических характеристик речи*, Издательство «Вебер», Донецк, 2008, 282 с.
- [6] Specht, D.F., “Probabilistic neural networks”, *Neural Networks, Vol. 3, 1990, p. 109-118*.
- [7] Бодянский, Е.В., Руденко, О.Г., *Искусственные нейронные сети: архитектуры, обучение, применения*, ТЕЛТЕХ, Харьков, 2004, 372 с.
- [8] Бондаренко, М.Ф., Дрюченко, О.Я., Коряк, С.Ф., Шабанов-Кушнарченко, Ю.П., *Идентификация людини за параметрами мовних сигналів (проблеми та шляхи їх вирішення)*, ТОВ «Компанія СМІТ», Харків, 2006, 260 с.
- [9] Атал, Б.С., “Автоматическое опознавание дикторов по голосам”, *ТИИЭР, Т. 64, №4, 1976, с. 48-66*.
- [10] Mitchell, M., *An introduction to genetic algorithms*, A Bradford Book The MIT Press, London, 1999, 158 p.
- [11] Скобцов, Ю.А., *Основы эволюционных вычислений: Учебное пособие*, ДонНТУ, Донецк, 2008, 326 с.