

Структурно-Параметрична Ідентифікація Правил Бази Модельних Знань

Ірина Жирякова

кафедра інтелектуальних систем прийняття рішень
Черкаський національний університет
імені Богдана Хмельницького
Черкаси, Україна
irena_zh@ukr.net

Сергій Голуб

кафедра інтелектуальних систем прийняття рішень
Черкаський національний університет
імені Богдана Хмельницького
Черкаси, Україна
fpkpk@ukr.net

Structural and Parametric Identification of the Knowledge Model Base Rules

Iryna Zhyriakova

Department of intelligent decision support systems
Bohdan Khmelnytsky National University
at Cherkasy
Cherkasy, Ukraine
irena_zh@ukr.net

Serhii Golub

Department of intelligent decision support systems
Bohdan Khmelnytsky National University
at Cherkasy
Cherkasy, Ukraine
fpkpk@ukr.net

Анотація—Обґрунтовано застосування нового підходу до автоматичного формування правил бази модельних знань для інтелектуальних систем прийняття рішень. Запропонований підхід базується на використанні основних положень теорії ієрархічних багаторівневих систем та методу багаторівневого моделювання складних систем за даними спостережень в умовах неповноти інформації.

Abstract—A new approach to automatic generation of the knowledge model base rules for intelligent decision support systems are proved. New approach is based on the main items of the hierarchical multilevel system theory and method multilevel modeling for complex systems.

Ключові слова—база модельних знань, інтелектуальні системи прийняття рішень

Keywords—knowledge model base, intelligent decision support systems

I. ВСТУП

Основні напрями і тенденції автоматизації інтелектуальної діяльності людини на основі використання сучасних інформаційних технологій досить різноманітні. Та прогрес в цій галузі значною мірою залежить від успішності використання процедур моніторингу для ідентифікації стану об'єктів інтелектуальної діяльності та їх взаємодії в рамках поставленої задачі. Такі об'єкти, зазвичай, являють собою складні системи, для

ідентифікації стану яких необхідно розв'язати комплекс задач в умовах невизначеності, адже, схильність до змін є невід'ємною характеристикою будь-якої предметної області.

Модель взаємодії складних систем в рамках будь-якої предметної області є основою процедурних знань (правил виводу) будь-якої інтелектуальної системи прийняття рішень. Отже, в даній роботі для ідентифікації правил виводу пропонується скористатись положеннями теорії ієрархічних багаторівневих систем [1], що дають можливість розглядати предметну область як декомпозицію етапів перетворення інформації, в результаті якої синтезуються багатопараметричні моделі, що можуть бути застосовані для представлення процедурної складової бази модельних знань. Переваги застосування ієрархічного підходу пов'язані зі зменшенням невизначеності стосовно як досліджуваних об'єктів так і середовища, в якому вони існують, на кожному етапі перетворення інформації.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ ТА ПОСТАНОВКА ПРОБЛЕМИ

Спробу повної систематизації алгоритмів і способів ідентифікації виконав один з основоположників теорії ідентифікації професор П. Ейкхоф [2]. Дана теорія є актуальною і по теперішній час при вивченні проблем побудови і всебічного аналізу моделей процесів або

систем. Вона описує поняття моделі, постановку задачі ідентифікації, базові підходи до побудови моделей для різних класів об'єктів, способи їх опису при різних підходах і алгоритмах ідентифікації.

Моделі подання знань почали розроблятися з 70-х років ХХ століття. Запропоновані підходи мали власну специфіку, яка найбільш помітна на синтаксичному рівні формального інструментарію для опису семантики. Застосування цих підходів досить широко використовувалось в діалогових системах типу «питання-відповідь» [3] та природно-мовних діалогових системах [4], а також в інтелектуальних розв'язувачах задач та системах управління [5-7]. Згодом проблематика даного напрямку досліджень, яка розглядалась в рамках штучного інтелекту, отримала назву "системи подання знань" [8]. До найбільш відомих моделей подання знань відносяться: семантичні мережі, концептуальні мережі, ситуаційне управління, формальні логічні моделі подання знань, універсальний семантичний код, мова Ф, фреймові моделі подання знань, зростаючі пірамідальні мережі, індикативні структури та онтології для опису семантики.

Огляд та аналіз існуючих моделей подання знань дає підстави стверджувати, що їх можливості обмежені в рамках вирішення задач оптимізації затрат часу на розробку бази знань або зміну її структури (підтримку актуальності у відповідності до нових задач). Таким чином, існує протиріччя між необхідністю розв'язку комплексу задач, що стоять перед сучасними інтелектуальними системами та обмеженими можливостями науково-методичного апарату, що дозволяє формалізувати роботу зі знаннями в автоматичному режимі. Наявність цього протиріччя з одного боку та існуюча і досі практика по формуванню баз знань експертами вручну або з використанням редакторів знань типу Protégé – з іншого боку, робить дослідження нових підходів до роботи з знаннями, в тому числі ідентифікацію процедурних знань, актуальними.

III. РЕЗУЛЬТАТИ

Згідно з описаною в [9, 10] загальною концепцією розробки бази модельних знань (БМЗ), що використовує методологію перетворення інформації на основі глобальної функціональної залежності [11], де кожна модель об'єкта в свою чергу є ієрархічною структурою, що містить інші моделі цього об'єкта, отримані за завершеними алгоритмами [12] для індуктивної ідентифікації правил БМЗ на основі емпіричних даних (навчаючих виборок, які формуються на основі масиву вхідних даних (МВД)) необхідно виконати спочатку структурну ідентифікацію досліджуваних об'єктів системи, а потім застосувати алгоритми параметричної ідентифікації. Розглянемо цей процес більш докладно на прикладі інтелектуалізації моніторингу «розумного міста» (Smart City).

Основною метою управління Smart City в рамках будь-якого міста є підвищення ефективності моніторингу всіх міських служб шляхом застосування інформаційно-комунікаційних технологій, тим самим усуваючи надмірність в генерації і використанні інформації. Отже, в межах окремих ланок управління інформаційні моделі

будуть являти собою локальні функціональні залежності (транспортна система, медицина, тощо), які поєднуються в єдину глобальну функціональну залежність в рамках обраного в якості об'єкта дослідження міста. Якщо ж розглядати подібну концепцію в масштабах регіону, то в процесі моделювання ми отримаємо вже цілий комплекс глобальних залежностей, які теж можуть бути поєднані в рамках системи управління країною. Отже, задача ідентифікації полягатиме у формуванні за даними моніторингу міських служб деякої множини моделей різної структури:

$$W = (X:Y), \quad (1)$$

де W – множина відображень показників стану модельованого об'єкта або змінні моделювання – вектор X , на модельований показник (вихідний сигнал), який потрібно визначити – Y .

Синтез будь-якого формального представлення моделі потребує рішення двох взаємозалежних задач: структурної ідентифікації і параметричної ідентифікації. Структурна ідентифікація передбачає визначення значущих факторів (змінних моделювання), що впливають на вихідні дані моделі (модельований показник); визначення структури, тобто виду оператора, що встановлює зв'язок між змінними моделювання і вихідними характеристиками моделі.

На першому етапі структурної ідентифікації аналізуються:

змінні моделювання, які утворюють топологічний простір X розмірності $\dim X = n \times m$, де n – кількість спостережень, де m – кількість змінних моделювання;

модельований показник Y розмірності $\dim Y = n \times l$, де n – кількість спостережень;

на основі топологічного простору формується система лінійно-незалежних векторів, що утворюватимуть матрицю МВД рангу $\text{rank} X = k, k \leq m$.

В процесі формування МВД дуже важливе значення мають методи стійкого оцінювання, що враховують наявність відхилень в даних і далі на етапі параметричної ідентифікації дозволяють їх компенсувати і досить точно визначати оцінки параметрів. Такі методи, називаються робастними або стійкими. Основи робастних методів оцінки були розроблені А.Н. Колмогоровим, Н.В. Смірновим і Б.С. Ястремським, а подальший розвиток отримали в роботах американських і швейцарських математиків [13]. В даній роботі для оцінки МВД пропонується застосування L- та E- критеріїв, запропонованих Г. Трітьєном та І.Г. Муром.

L-критерій має два різновиди застосування: для виявлення грубих помилок у верхній частині L та у нижній частині L' ранжованого ряду даних:

$$L = \frac{\sum_{i=1}^{n-l} (x_i - \bar{x}_l)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}, L' = \frac{\sum_{i=l+1}^{n-l} (x_i - \bar{x}_l)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (2)$$

де x_i – вибірка даних спостережень по одній обраній j -тій ознаці, n – обсяг вибірки, l – кількість спостережень з відхиленнями в значенні обраної ознаки, \bar{x} – загальна середня для вибіркової сукупності, \bar{x}_l – середнє значення для $(n-l)$ спостережень, що залишились після відкидання l грубих помилок у верхній для L (3) та у нижній частині – для L' (4) ранжованого ряду даних.

$$\bar{x}_l = \frac{\sum_{i=1}^{n-l} x_i}{n-l} \quad (3)$$

$$\bar{x}_l = \frac{\sum_{i=l+1}^n x_i}{n-l} \quad (4)$$

E -критерій застосовується у випадку, коли у вибірці грубі помилки з найбільшими і найменшими значеннями розташовані як у верхній, так і у нижній частині ранжованого ряду даних:

$$E = \frac{\sum_{i=1}^{n-l} (x_i - \bar{x}_l)^2}{\sum_{i=l+1}^n (x_i - \bar{x})^2}, \quad (5)$$

де \bar{x}_l – середнє значення після відкидання найменших l і найбільших l' грубих помилок, яка обчислюється за формулою:

$$\bar{x}_l = \frac{\sum_{i=l+1}^{n-l'} x_i}{n-(l+l')}. \quad (6)$$

Всі три критерії мають табличні критичні значення для заданого рівня значущості α при відомому обсязі вибірки n і передбачуваній кількості помилок l . Якщо значення критеріїв виявляються менше порогових, то помилки в даних, визнаються грубими. Інакше дані вважаються типовими для даної сукупності спостережень.

На другому етапі структурної ідентифікації для вибору класу моделей можна скористатись будь-якою з двох аксіоматичних структур: адитивною (7) або мультиплікативною (8), що визначають специфікацію моделі в математичній формі. Відповідний вибір буде цілком залежати від підсистеми управління для вибраного об'єкту моніторингу. Подібні аксіоматичні структури у

вигляді поліномів і різницевих рівнянь, які відносяться до класу моделей лінійних за параметрами, зазвичай використовують для моделювання статичних або динамічних об'єктів.

$$P(x) = \sum_{i=1}^n a_i x_i, \quad (7)$$

$$P(x) = \prod_{i=1}^n a_i x_i, \quad (8)$$

де a_i – відносні безрозмірні вагові коефіцієнти, для яких виконуються наступні обмеження:

$$0 \leq a_i \leq 1, \sum_{i=1}^n a_i = 1. \quad (9)$$

Недоліками адитивної структури є відсутність врахування наявності можливої нелінійності зв'язку. Цей недолік можна усунути при більш широких обмежуючих допущеннях про структуру моделі, наприклад, скориставшись поліномом Колмогорова-Габора:

$$P(x) = a_0 + \sum_{i=1}^n a_i x_i + \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j + \dots \quad (10)$$

Такий підхід дозволяє описати будь-яку нелінійну залежність і не накладає ніяких апріорних обмежень на адитивність або мультиплікативність, тому що поліном (10) містить у своєму складі як перші, так і більш високі ступені змінних моделювання і всі можливі їх комбінації.

Отже, обравши відповідну аксіоматичну структуру для специфікації моделі задача її синтезу зводиться до параметричної ідентифікації вагових коефіцієнтів a_i .

Алгоритм синтезу моделі (АСМ), що передбачає визначення конкретних кількісних значень її параметрів, є найбільш формалізованим етапом і, як правило, зводиться до вирішення задачі оптимізації із мінімізації середньої квадратичної нев'язки для залежності з неперервним виходом чи рівня безпомилковості класифікації для залежності з дискретним виходом.

Для реалізації АСМ можна використати перебірний (комбінаторний) або ітераційний (багаторядний) метод генерації структури моделі (11). Застосування будь-якого з них, дасть можливість сформувати множину моделей-кандидатів (12), серед яких буде обрана оптимальна за визначеним критерієм якості (13):

$$\hat{Y}_f = f(X, \hat{\theta}_f), f \in \Phi, \quad (11)$$

де Φ – множина моделей-кандидатів, $\hat{\theta}_f$ – оцінка параметрів для кожної f .

$$f^* = \arg \min_{f \in \Phi} C(y, f(X, \hat{\theta}_f)), \quad (12)$$

де C – критерій якості моделей, f^* – оптимальна модель серед усіх $f \in \Phi$.

$$\hat{\theta}_f = \arg \min_{\theta \in R^s} Q(y, X, \theta_f), \quad (13)$$

де $Q \neq C$ – критерій якості оцінювання параметрів кожної окремої моделі при розв'язанні задачі параметричної ідентифікації, що генерується в процесі структурної ідентифікації.

Можлива різна композиція методів структурної і параметричної ідентифікації, яка буде відрізнятися лише точністю, трудомісткістю і універсальністю.

Відповідно до методу висхідного синтезу елементів, кожна глобальна функціональна залежність буде ієрархічним поєднанням локальних багатопараметричних моделей, згенерованих відповідним АСМ, та поєднаних в страти відповідно до рівнів моніторингу:

$$Z = f(Y(X)). \quad (14)$$

Така структура забезпечуватиме багаторівневе перетворення інформації від форми чисельних характеристик результатів спостереження до форми відомостей про властивості об'єктів, поданих в зручному для використання вигляді.

Кожна із страт матиме декілька недостатньо інформативних моделей, вихідні сигнали яких мають характеристики гірше заданих. Це пов'язано з тим, що задача параметричної ідентифікації є некоректною по Адамару, так як не має єдиного розв'язку. Згідно з теоремою Вейерштрасса, будь-яку послідовність з n експериментальних даних можливо точно апроксимувати поліномом із $(n + 1)$ членом. Однак, така апроксимація не означає, що синтезовано адекватну модель високої точності з гарними прогностичними властивостями. Отже, будь-який поліном високої складності апроксимує не тільки корисний сигнал, але й випадкові похибки експериментальних даних, які частково усуваються L- та E- критеріями.

Для подолання зазначеного недоліку в роботі [14] запропоновано розділяти вибірку експериментальних даних на дві підмножини: навчальну і перевірочну. Перша підмножина використовується для синтезу моделі і визначення її характеристик, а друга – для перевірки точності моделі. При цьому виявилось, що при збільшенні складності моделі точність апроксимації перевірочної послідовності експериментальних даних спочатку поліпшується, досягає деякого мінімуму і потім починає погіршуватися за рахунок обліку “шкідливих” випадкових складових. Модель мінімальної складності, що дає мінімум похибки апроксимації перевірочної послідовності, одержала назву моделі оптимальної складності.

Періодичне поновлення та доповнення БМЗ останніми результатами моніторингу та автоматичне адаптування структури БМЗ за описаними правилами дозволятиме врахувати динамічність процесів управління будь-якою ланкою Smart Sity та забезпечить інформацією процеси прийняття управлінських рішень на кожній з них.

IV. ВИСНОВКИ

У даній роботі запропоновано новий підхід до структурно-параметричної ідентифікації правил для баз модельних знань, який базується на наступних принципах: знання представляються у вигляді багатопараметричних моделей; зростання кількості знань досягається шляхом ієрархічного поєднання багатопараметричних моделей відповідно методу висхідного синтезу елементів в системах з багаторівневим перетворенням даних; використання положень теорії ієрархічних багаторівневих систем є основою для розв'язання задач структурної ідентифікації; консолідація знань відбувається шляхом поєднання глобальних функціональних залежностей, сформованих в рамках різнорідних предметних областей. Запропонований підхід орієнтований на автоматичне формування та подальший автоматизований супровід баз знань для будь-якої предметної області.

ЛІТЕРАТУРА REFERENCES

- [1] М. Месарович, Теория иерархических многоуровневых систем / М. Месарович, Д. Мако, И. Тахакара. – М.: Мир, 1973. – 344 с.
- [2] П. Эйкхофф, Основы идентификации систем управления. – М.: Мир, 1975. – 686 с.
- [3] Г.Г. Белоголов, Б.А. Кузнецов, Языковые средства автоматизированных информационных систем. – М.: Наука, 1983. – 288 с.
- [4] Э.В. Попов, Искусственный интеллект: Кн.1. Системы общения и экспертные системы: справочник / под ред. Э.В. Попова. – М.: Радио и связь, 1990. – 464 с.
- [5] Н. Нильсон, Принципы искусственного интеллекта. – М.: Радио и связь, 1985. – 372 с.
- [6] Д.А. Поспелов, Логико-лингвистические модели в системах управления. – М.: Энергоиздат, 1981. – 231 с.
- [7] М. Минский, Фреймы для представления знаний. – М.: Энергия, 1979. – 151 с.
- [8] Т.А. Гаврилова, В.Ф. Хорошевский, Базы знаний интеллектуальных систем. – СПб.: Питер, 2001 – 384 с.
- [9] І.А. Жирякова, Новий підхід до концептуалізації знань / І.А. Жирякова, С.В. Голуб // Технічні науки та технології. – 2015. – № 2. – С. 78-82.
- [10] І.А. Жирякова, Удосконалення методу висхідного синтезу елементів в інформаційній технології багаторівневого моніторингу мобільного робота / І.А. Жирякова, С.В. Голуб, О.В. Харченко // Математичні машини і системи. – 2016. – № 3. – С. 41-47.
- [11] С.В. Голуб, Багаторівневе моделювання в технологіях моніторингу оточуючого середовища. – Черкаси: Вид. від. ЧНУ імені Богдана Хмельницького, 2007. – 220 с.
- [12] С.В. Голуб, Координація взаємодій локальних агрегатів в структурі систем багаторівневого перетворення моніторингової інформації // Вісник Східноукраїнського національного університету імені Володимира Даля. – 2009. – № 6(136). – Ч. 1. – С. 325-329.
- [13] А.М. Шурыгин, Прикладная стохастика: робастность, оценивание, прогноз. – М.: Финансы и статистика, 2000. – 224 с.
- [14] В.С. Степашко, Структурна ідентифікація моделей як задача відновлення сигналу в умовах неповноти інформації // Наукові праці ДНТУ. Серія: Обчисл. техніка та автоматизація. – Вип. 48. – Донецьк: ДНТУ, 2002. – С. 345-353.