Проблема інтелектуалізації прийняття рішень у задачах ідентифікації моделей

В.С.Степашко, Т.Ф.Зворигіна, С.М.Єфіменко
Міжнародний науково-навчальний центр інформаційних технологій та систем НАН та МОН України, Київ, проспект Глушкова, 40, astrid@irtc.org.ua

Вступ. Прийняття рішень у різних сферах людської діяльності вимагає наявності засобів, що допомагають отримувати точні моделі для прогнозування процесів. Складні проблеми, що виникають при моделюванні, пов'язані з великим числом змінних, малою кількістю спостережень і невідомими динамічними зв'язками між змінними. Такі проблеми характеризуються: недостатньою апріорною інформацією; великою кількістю параметрів; зашумленими або короткими вибірками даних.

Нині існує багато різних методів моделювання за даними спостережень (МДС), або структурно-параметричної ідентифікації (СПІ) [1], однак апріорно невідомо, який з методів дасть найкращий результат у конкретній задачі, а ефективність деяких з них є недостатньо дослідженою. Успішне розв'язання задачі моделювання складних об'єктів, процесів і систем за даними спостережень значною мірою залежить від знання особливостей цих методів.

Отже, для вирішення проблеми інтелектуалізації необхідно розв'язати дві взаємопов'язані задачі: розроблення імітаційного комплексу для проведення випробувань методів моделювання за даними спостережень, а також створення інтелектуальної оболонки, яка в доступному для користувача будь-якого рівня підготовленості режимі допоможе пройти весь шлях розв'язання задачі МДС – від первинної обробки таблиці даних до формування алгоритму моделювання та його застосування.

Такий імітаційний комплекс дає можливість:

- розробляти методики і планувати статистичні випробування;
- тестувати різні методи моделювання та їхні компоненти;
- порівнювати наявні методи за певними критеріями;
- конструювати методи моделювання за даними спостережень;
- проводити імітаційні експерименти (екстраполяцію, прогнозування) з моделями, побудованими за різними методами моделювання;
- поповнювати свої знання про методи моделювання під час роботи з імітаційним комплексом.

В основі роботи комплексу лежить метод Монте-Карло (метод статистичного експерименту) як гнучкий і потужний засіб дослідження та оцінки ефективності методів моделювання.

Виходячи з вимог до імітаційного комплексу, зазначимо, що це має бути досить складний програмний продукт, створення якого неможливе без

ґрунтовних знань предметної області, тому далі зупинимось на особливостях методів моделювання за даними спостережень.

Аналіз структури процесу і методів моделювання. У загальному випадку процес розв'язання задачі моделювання за даними спостережень складається з таких основних етапів [2]:

1. Задання вибірки даних (отриманої в результаті пасивного або активного експерименту), а також апріорної інформації;
2. Вибір або задання класу базисних функцій і відповідного перетворювання даних;
3. Генерація різних структур моделей у вибраному класі;
4. Оцінювання параметрів структур, що генеруються, і формування множини моделей;
5. Вибір критерію якості кожної зґенерованої моделі;
6. Мінімізація заданого критерію і вибір оптимальної моделі f;
7. Перевірка адекватності отриманої оптимальної моделі;
8. Застосування моделі (її використання).

Кожен метод моделювання можна описати за допомогою чотирьох компонентів, за якими можна порівнювати функціональні особливості заданої множини методів [2]. З іншого боку, для певної множини методів, розбивши кожен з них на такі складові, можна утворити множину класів моделей, генераторів структур, методів оцінювання параметрів і критеріїв відбору моделей і аналізувати ці множини як самостійні об'єкти. Таке розбиття дає можливість сформувати функціональне наповнення програмної системи.

Інструментальні засоби моделювання. Розробляється інструментальний комплекс взаємопов'язаних програм, що охоплюють різні функціональні варіанти вказаних чотирьох компонентів. Це буде імітаційний комплекс програм, що має модульну структуру.

У роботі представлено результати тестових експериментів, проведених за допомогою готових модулів такого програмного комплексу:

- порівняння швидкодії параметричної ідентифікації з застосуванням чотирьох методів розв'язання систем лінійних рівнянь [3]: Краута, Гауса, методу ортогоналізації Грама-Шмідта та рекурентного методу обрамлення;
- порівняння швидкодії структурної ідентифікації з використанням відомих та нових рекурентних методів оцінювання параметрів [4];
- порівняння ефективності використання деяких генераторів структур: клас вкладених структур, комбінаторний алгоритм, метод включення та комбінаторно-селекційний алгоритм;
- порівняння ефективності використання деяких критеріїв структурної ідентифікації моделей.

Результати проведених тестових випробувань дають можливість сформулювати рекомендації щодо порівняльної ефективності застосування тих чи інших алгоритмів структурно-параметричної ідентифікації.

Інтелектуальна підтримка процесу ідентифікації. Другою частиною вирішення проблеми інтелектуалізації є створення інтелектуальної оболонки.

При її конструюванні необхідно враховувати як побажання користувача і невизначеності, про які йшлося у вступі, так і інформацію, що неявно міститься у вибірці даних. Тобто для побудови такої оболонки необхідно розв'язати задачу автоматизації процесу побудови найкращого алгоритму. Можливість такої автоматизації з'являється за рахунок використання концепції знання-орієнтованого конструювання алгоритму СПІ.

В [4] було сформульовано основні вимоги до такої оболонки:

- реалізація експертних знань;
- інтерактивна підтримка всіх етапів моделювання;
- мінімізація вимог до рівня підготовки користувача;
- активне використання знань користувача;
- постійний контроль і перевірка несуперечливості рішень, що приймаються користувачем;
- наочність процесу розв'язання і контекстність сприйняття інформації;
- обов'язкове розв'язання задачі від початку до кінця;
- прийняття рішення у разі відмови користувача (процедура „не знаю");
- несуперечливість автоматичних процедур, цілісність бази знань;
- відкритість системи, тобто можливість наповнення новими методами;
- навчання користувача в процесі роботи з системою.

Очевидно, що для створення системи з дружнім інтерфейсом вказаного вище переліку етапів недостатньо. Користувач, а ним в нашому випадку буде фахівець-прикладник, може мати багато апріорної інформації про суть задачі, але не зобов'язаний уміти проводити кореляційний аналіз чи визначати порядок авторегресійної складової моделі.

Тому було сформульовано таку модифіковану послідовність етапів прийняття рішень, що задовольняє вказані вимоги:

1. Визначення мети моделювання.
2. Визначення типу модельованого процесу.
3. Виявлення стаціонарності процесу або виду тренду.
4. Визначення класу (класів) моделей.
5. Оцінка типу розподілу шуму.
6. Визначення критерію відбору (селекції) кращих моделей.
7. Визначення методу оцінювання параметрів.
8. Вибір генератора структур (організатора перебору) моделей.
9. Реалізація алгоритму.

Специфіка такої послідовності полягає в тому, що кожен етап складається зі скінченої множини можливих варіантів рішень; на кожному етапі користувач вибирає якесь одне опорне рішення, і тоді повна множина рішень на кожному з подальших етапів розпадається на припустимі і неприпустимі підмножини, за рахунок чого кількість припустимих рішень зменшується [6]. Таким чином, на

етапі формування (реалізації) алгоритму ми маємо, взагалі кажучи, всього кілька доцільних варіантів рішень, а в ідеалі – єдиний варіант.

Це досягається завдяки такій впорядкованості етапів, що кожен наступний етап залежить або від попереднього, або комбінації варіантів на попередніх етапах, і не залежить від подальших етапів.

Висновок. Моделювання за даними спостережень, або структурно-параметрична ідентифікація, за виразом Л.Льюнга, це скоріше мистецтво, ніж наука [7]. До мистецтва відносять уміння дослідника за інформацією, що неявно міститься у вибірці даних, а також на основі апріорних знань сформулювати вимоги до моделі, адекватної меті моделювання. Науковою (формалізованою) є інша частина задачі – розробка безпосередньо алгоритму моделювання та побудова на його основі математичної моделі. Запропонований у доповіді підхід до послідовної побудови алгоритмів із заданої кількості їх конструктивних елементів дозволяє розширити наукову частину за рахунок виявлення і реалізації наявних експертних знань та „здорового глузду”.

Література

1. Степашко В.С., Кочерга Ю.Л. Методы и критерии решения задач структурной идентификации // Автоматика. – 1985. – №5. – С.29-37.
2. Степашко В.С. Алгоритмы МГУА как основа автоматизации процесса моделирования по экспериментальным данным // Там же. – 1988. – № 4. – С.44-55.
3. Райс Дж. Матричные вычисления и математическое обеспечение: Пер. с англ. – М.: Мир, 1984. – 264с.
4. Степашко В.С., Єфименко С.М. Аналіз особливостей рекурентного методу послідовного оцінювання параметрів // Відбір і обробка інформації. – 2004. – Вип. 21. – С. 91-96.
5. Зворыгина Т.Ф., Степашко В.С. Концепция знание-ориентированного конструирования алгоритма моделирования процессов и объектов по данным наблюдений // Матер. Міжнар. конф. з автоматичного управління „Автоматика-2004”. – С. 47.
6. Степашко В.С., Зворыгина Т.Ф. Об одном подходе к проблеме вывода решений в сложной задаче // УСиМ. – 2003. – № 6. – С.82-87.
7. Льюнг Л. Идентификация систем. – М.: Наука, 1991. – 431 с.