

# Застосування Багатомодельного Класифікатора в Інформаційних Системах Кризового Моніторингу

А. Авраменко

кафедра інтелектуальних систем прийняття рішень  
Черкаський національний університет  
імені Богдана Хмельницького  
Черкаси, Україна  
RedStar929@gmail.com

С. Голуб

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

## Application of Multimodel Classifier in the Information Systems for Crisis Monitoring

A. Avramenko

Department of intelligent decision support systems  
Bogdan Khmelnytsky Cherkassy National University  
Cherkassy, Ukraine  
RedStar929@gmail.com

S. Golub

Department of intelligent decision support systems  
Bogdan Khmelnytsky Cherkassy National University  
Cherkassy, Ukraine  
fpkpk@ukr.net

**Анотація**—Запропоновано покращення методу класифікації масивів вхідних даних в інформаційних системах кризового моніторингу. Зменшення часу оптимізації досягнуто шляхом використання багатомодельних класифікаторів. Класифікація проводиться за допомогою групового рішення сукупності класифікаторів. Експериментально підтверджено ефективність покращеного методу. Час перебудови структури зменшується в 2-3 рази, в порівнянні з не модернізованим методом. Похибка моделювання значимо не погіршується.

**Abstract**—Improvement of method for classification of the input data arrays in the crisis monitoring information systems is proposed. Reducing time of optimization achieved by using multimodel classifiers. Classification is conducted by the group decision of classifiers. Effectiveness of the improved method was experimentally confirmed. Time for restructuring of models reduced by 2-3 times compared to unimproved method. Errors of modelling is not significantly worse.

**Ключові слова**—кризовий моніторинг, багаторівневі класифікатори багаторівневе моделювання, час перебудови системи, похибка моделювання.

**Keywords**—crisis monitoring, multimodel classifiers, multilevel modeling, group decision, system restructuring time, modelling error

### I. ВСТУП

Використання та розробка моніторингових систем є надзвичайно важливим питанням як наукових досліджень, так і практичного застосування. Моніторингові системи дозволяють робити висновки про певні події, базуючись при цьому лише на попередні дані, отримані шляхом спостереження. Саме тому дані системи досить успішно використовують у якості систем діагностики у різних сферах життя.

Склад системи моніторингу визначається наступними чинниками: функціональне призначення, область застосування та цільові установки з переліку вирішуваних задач, функції обробки інформації, покладені на систему моніторингу і визначені користувачем.

Основним завданням моделювання в моніторингових системах є забезпечення інформацією процесу прийняття рішень. Інформація отримується в результаті моделювання властивостей об'єкта моніторингу на основі даних, отриманих в процесі вимірювання чисельних характеристик цього об'єкта.

Саме зараз такі системи використовуються і для моніторингу надзвичайних ситуацій. Проте прийняття рішень в умовах надзвичайних ситуацій накладає ряд обмежень на технології, що забезпечують ці процеси інформацією.

Розвиток надзвичайних ситуацій є ланцюговим лавиноподібним динамічним процесом. Він полягає в

різкому погіршенні стану сукупності територій, які можуть містити об'єкт чи об'єкти економіки і житлових комплексів, що призводить до катастрофічних для цих об'єктів та їх оточення наслідків.

Рішення в даних ситуаціях повинні отримуватися як найшвидше. Також незадовільна прогнозованість та динамічність надзвичайних ситуацій генерує велику кількість параметрів для моделювання, частина з яких може виникнути вперше, що означає велику можливість помилки в попередньо отриманих моделях.

На сьогодні одними з перспективних систем моніторингу є системи засновані на технології багаторівневого перетворення даних, яка реалізована у вигляді інформаційної системи з ієрархічним поєднанням багатопараметричних моделей [1].

Такі моделі можуть бути синтезовані за допомогою індуктивних алгоритмів, нейронних мереж, генетичних алгоритмів та інших. В даній технології цій системі сценарій вибору кращого алгоритму синтезу багатопараметричних моделей (АСМ) реалізовано шляхом послідовного їх випробування та вибору кращого [2].

З синтезованих моделей формується ієрархія. Моделі на кожному рівні ієрархії розв'язують локальні задачі із перетворення даних. В таких ієрархічних структурах можуть поєднуватися велика кількість моделей, від п'ятдесяти і більше (рис. 1).

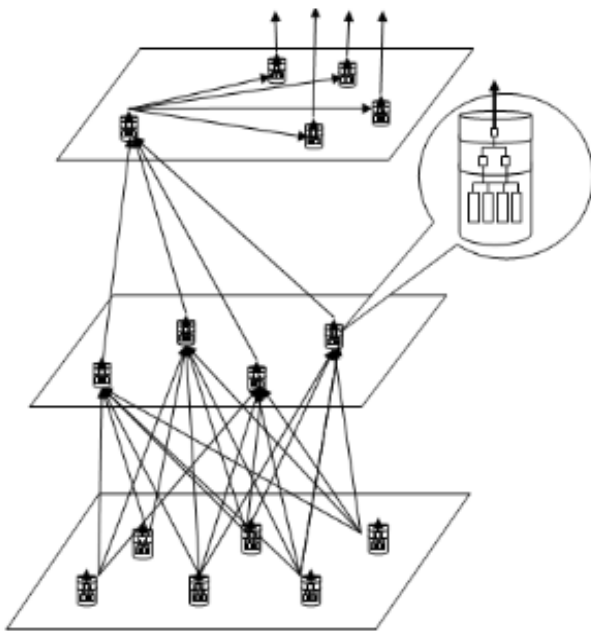


Рис. 1. Структура системи перетворення інформації

В процесі моніторингу кризових ситуацій властивості масиву вхідних даних постійно змінюються. Тому існує велика можливість, що одна або декілька моделей можуть почати видавати не адекватні результати. Для виправлення таких «пошкоджень» проводиться заміна цих моделей з їх повторним синтезом та синтезом усіх моделей, що з ними пов'язані.

Процес синтезу усіх «пошкоджених» моделей та моделей, що з ними пов'язані займає тривалий час, приблизно 40 хвилин. Тривалість залежить від кількості моделей в структурі.

Зважаючи на те, що в умовах кризового моніторингу на обґрунтування рішень виділяється не більше 2-3 хвилин, а властивості МВД змінюються динамічно. Тому є необхідність в зменшенні часу перенавчання ієрархічної системи моделей.

В попередній роботі [3] було запропоновано покращення існуючого алгоритму шляхом класифікації. А саме, на відміну від послідовного випробування кожного із наявних АСМ із наступним вибором кращого алгоритму, було запропоновано побудувати вирішальне правило, яке забезпечить для кожного МВД вибір найбільш придатного АСМ.

Вибір найбільш придатного АСМ проводиться на основі таких інформативних параметрів таблиць первинного опису [6]:

- кількість спостережень;
- кількість незалежних змінних;
- кількість параметрів максимально суміщених з функцією мети;
- кількість не суміщених параметрів;
- середній коефіцієнт кореляції незалежних змінних;
- середній коефіцієнт кореляції незалежних змінних та функції мети;
- середній коефіцієнт детермінації незалежних змінних;
- середній коефіцієнт детермінації незалежних змінних та функції мети;
- визначник нормованої таблиці первинного опису;
- визначник нормованої матриці значень незалежних змінних;
- власне число нормованої таблиці первинного опису;
- власне число матриці значень незалежних змінних.
- максимальне сингулярне число нормованої таблиці первинного опису;
- максимальне сингулярне число матриці значень незалежних змінних.

Дане дослідження ґрунтується на вивченні результатів попередньої роботи та аналізу сучасних методів, які дозволяють покращити результати роботи АСМ та навчання ієрархічних систем моніторингу.

За результатами дослідження методу запропонованого в [3], виявлено, що метод має певні недоліки, а саме - погіршення показників якості змодельованих моделей. Зважаючи на те, що в структурі інформаційної системи багаторівневого перетворення даних міститься від п'ятдесяти моделей і більше, вдається досягнути значного скорочення часу адаптації структури системи до зміни

властивостей МВД. Проте, у той самий час, ми не маємо можливості задалегідь виявляти, які моделі будуть мати погіршення або покращення в порівнянні з повним перебором.

Середній результат моделювання всієї ієрархічної структури при використанні методу класифікації може навіть, у деяких випадках, бути кращим за повний перебір, а у випадку втрат точності було запропоновано компенсувати їх за допомогою збільшення складності багатопланової ієрархічної структури системи. Це досягається завдяки ефекту емерджентності властивостей багаторівневих ієрархічних систем.

Отже результати моделювання кожної моделі є не передбачуваними і можуть як поліпшитись так і погіршитись, в порівнянні з повним перебором у стандартній реалізації.

Об'єктом нашого моніторингу є кризові ситуації, це накладає ряд обмежень на отримані моделі. Зокрема, для правильної інтерпретації результатів ми повинні мати довіру до результатів роботи кожної отриманої моделі, а не тільки до кінцевого результату. А так як метод запропонований у [3] отримує свої результати при не передбачуваних втраті точності моделей, тому з'являється необхідність його покращення з метою зменшення втрат точності.

Було розглянуто метод що дозволяє покращити результати моделювання шляхом використання особливостей формування МВД у багаторівневих ієрархічних системах [4].

Даний метод створює дублікати рівнів у ієрархії, що при особливостях утворення МВД дозволяє покращувати результати вже змодельованих моделей створивши багатопланову систему з дублікатів цих моделей.

## II. МЕТА РОБОТИ

Метою цієї роботи є зменшення часу перебудови структури ієрархії моделей без значної втрати точності результатів моделювання на виході системи.

Для досягнення поставленої мети було запропоновано вдосконалити метод сформульований у попередній роботі [3] та усунути його недоліки.

При вирішенні даної задачі було сформульовано наступні гіпотези:

Спостереження, з якими класифікатор не справився можливо використати для навчання іншого класифікатора, що буде впізнавати саме ці дані

Можливо отримати кілька класифікаторів, які виконують класифікацію одного і того самого АСМ

Набір класифікаторів що виконують класифікацію МВД для одного і того ж АСМ можливо використовувати як експертів у питанні класифікації для даного АСМ, а результати поєднувати за допомогою методів прийняття групових рішень [5]

Моделі класифікаторів можливо поліпшити використавши метод адаптивного дублювання рівнів [4]

## III. РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ

При поєднанні висловлених гіпотез було сформульовано наступний алгоритм навчання класифікаторів.

Після проведення навчання класифікаторів на етапі тестування ми отримуємо новий МВД для навчання наступного класифікатора, якщо це можливо. Навчаємо його, а результати їх роботи поєднуємо за допомогою методу простого голосування [5].

Далі проводиться тестування поєднаного рішення класифікаторів та отримується новий МВД, для якого поєднаний результат не розпізнав найкращий АСМ. Всі ці дії повторюються до отримання нового МВД, за допомогою якого вже неможливо навчити наступний класифікатор. Результатом виконання даного алгоритму є отримання сукупності моделей та особливості їх групової взаємодії.

Кожна модель отримана за даним алгоритмом є рівноправною, а отже ієрархічно вони знаходяться на одному рівні. Тому ці моделі можливо подальше покращити за допомогою методу адаптивного дублювання рівнів [4].

Для проведення дослідження було модернізовано програмний інструмент створений у [3], в якому реалізовані запропоновані покращення.

Метою дослідження було:

- Перевірка доцільності запропонованих покращень
- Подальше тестування роботи методу класифікації

Для синтезу моделей даного дослідження використані результати моніторингу захворюваності населення Черкаської області впродовж 2000-2014 років [1].

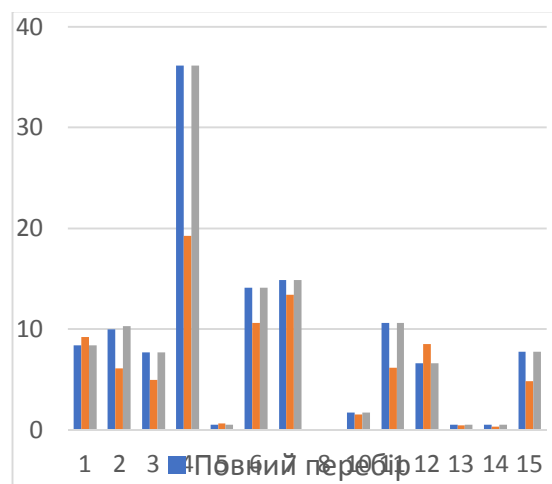


Рис. 2. Порівняння СКО

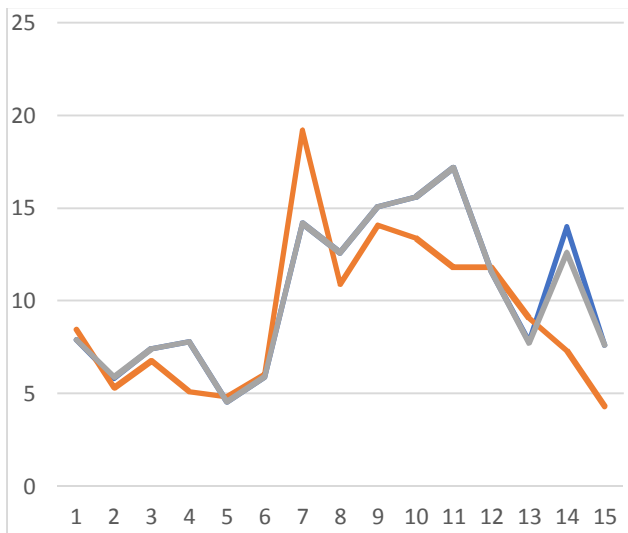


Рис. 3. Порівняння абсолютних похибок

На даних діаграмах рис.2 та рис.3 зображено відповідно порівняння стандартного та абсолютного відхилення між алгоритмом повного перебору, алгоритмом з використанням одномодельного класифікатора та запропонованим у даній роботі багатомодельним варіантом класифікатора.

В результаті можна побачити, що у деяких випадках одномодельний класифікатор має кращі результати, а у деяких гірші, ніж повний перебір АСМ. Але в середньому одномодельний класифікатор має кращий результат. Якщо порівнювати з багатомодельним варіантом, то він має результат що майже повністю сходиться з алгоритмом повного перебору.

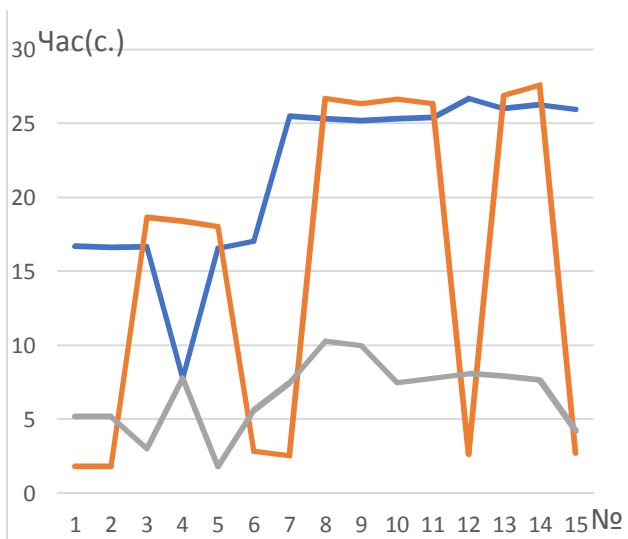


Рис. 4. Порівняння часу навчання

При порівнянні часу навчання рис.4 можливо побачити як саме нестабільно працює одномодельний

класифікатор. Покращений метод в середньому працює швидше, а отже призводить до швидшого навчання системи.

#### IV. ВИСНОВКИ

В результаті порівняння роботи повного перебору, методу із [3] та його покращений багатомодельний варіант виявилося, що запропонований метод майже сходиться з повним перебором по таким критеріям як абсолютна та стандартна похибка, різниця між похибками не перевищує 1%. В той самий час при використанні звичайного одномодельного класифікатора різниця відхилень сягає 30-40%. Схожі результати було отримано і при порівнянні часу навчання ієрархічних систем.

В середньому час навчання системи з використанням нового покращеного класифікатора зменшився в два рази при незначних втратах якості моделей. Зважаючи на те, що якість моделей гірша тільки в порівнянні з методом запропонованим у [3] можливо стверджувати, що поставлена мета дослідження була досягнута.

Дані результати можуть свідчити про досягнення системності при використанні багатомодельних класифікаторів.

Отже, за даними дослідження запропоновані покращення методу класифікацій алгоритмів синтезу моделей дійсно дозволяють усунути знайдені недоліки та демонструють можливість подальшого вдосконалення що потребує вивчення.

#### ЛІТЕРАТУРА REFERENCES

- [1] Голуб С.В. Багаторівневе моделювання в технологіях моніторингу оточуючого середовища / С.В. Голуб. – Черкаси: Вид. від ЧНУ імені Богдана Хмельницького, 2007. – 218 с.
- [2] Голуб С.В. Застосування стратегії оптимальності при виборі алгоритмів синтезу моделей у системах багаторівневого соціоecологічного моніторингу / С.В. Голуб, П. О. Колос // Математичні машини і системи. - 2010. - № 4. - С. 127 - 134.
- [3] Avramenko A.: Classification models in information systems for social and environmental crisis monitoring / Avramenko A., Golub S. // Engineer of XXI Century – We design the future: Wydawnictwo Naukowe Akademii Techniczno-Humanistycznej w Bielsku-Bialej - Bielsku-Biala : ATH, 2016. – 928 p.
- [4] Адаптивне формування дублюючих рівнів в структурі ієрархічних систем багаторівневого соціогігієнічного моніторингу / С.В. Голуб, В.Ю. Немченко // Індуктивне моделювання складних систем: Зб. наук. пр. — К.: МННЦ ІТС НАН та МОН України, 2011. — Вип. 3. — С. 41-48. — Бібліогр.: 8 назв. — укр.
- [5] Барабаш Ю.Л. Коллективные статистические решения при распознавании. — М.: Радио и связь, 1983, —224 с., пл.
- [6] Колос П.О. Визначення множини інформативних параметрів таблиці первинного опису об'єкта моделювання./ Вісник Черкаського університету, випуск 173. – Черкаси: Вид. ЧНУ, 2009. – С. 121-128.