

## **COMPLEX SYSTEMS STATE OVERALL ESTIMATION SYNTHESIS BY CLUSTERING METHODS**

**S. Chumachenko, A. Moshensky**

*National University of Food Technologies*

**I. Savitskaya, V. Smolii**

*National University of Life and Environmental Sciences of Ukraine*

---

**Key words:**

*Cluster*  
*Soft-clustering*  
*System state*  
*Image*  
*Integral estimation*

**ABSTRACT**

The object of research is the coding process of the current system overall state assessment.

The analysis of complex systems characteristics indicates systems with uncertain final state in the control process, since there are several valid options for its implementation. It was also established that the entire system and its components' states can be set by discrete integer values or brought to this form. The total number of possible states is small and pre-determined. These features indicate difficulties with methods for determining the integral estimation of the system state, since they do not allow to single out its individual components. The use of pattern recognition methods is quite complex and resource intensive.

The goal of the work is synthesis methods development for a general estimation of the complex system state with ability to control its components state. The subject of the study is codes with a separate representation of the components.

To solve this problem, we proposed cluster analysis methods that, in the applied sense, set the goal of determining the cluster number of the control object. The considered features of representing the system state and its components indicate the possibility of using soft clustering methods with grouping.

In the work, the synthesis method for a general estimation of the complex system state is proposed and considered in terms of a coal-mining combine speed control system. The use of clustering allowed the individual clusters allocation in the set of system states and the principles of grouping estimates of subsystems' state.

The conclusions of the work substantiate a high level of synthesized code adaptation for computer networks and systems.

---

**Article history:**

Received 20.04.2020

Received in revised form

19.05.2020

Accepted 28.05.2020

---

**Corresponding author:**

S. Chumachenko

**E-mail:**

sergiy23.chumachenko@gmail.com

## СИНТЕЗ ОЦІНКИ СТАНУ СКЛАДНОЇ СИСТЕМИ МЕТОДАМИ КЛАСТЕРИЗАЦІЇ

С. М. Чумаченко, А. О. Мошенський

Національний університет харчових технологій

Я. А. Савицька, В. В. Смолій

Національний університет біоресурсів і природокористування України

*У статті досліджено процес кодування сукупної оцінки стану складної системи. Проведений аналіз характеристик складних систем показує, що існує ряд систем, кінцевий стан яких у процесі керування невизначений, оскільки є декілька допустимих варіантів його реалізації. Крім того, з'ясовано, що система цілком та її окремі складові знаходяться у станах, які можна задати дискретними цілочисельними значеннями або привести їх до такого виду. Загальна кількість можливих станів є невеликою та визначеною. При визначенні керуючого впливу на систему важливо також знати, який стан має кожна з складових цієї системи.*

*Ці особливості вказують на труднощі використання методів визначення інтегральної оцінки стану системи, оскільки вони не дають змоги виділити окремі її складові. Використання методів розпізнавання образів є досить складним та ресурсоємним, тому мета дослідження передбачає розробку методів синтезу сукупної оцінки стану складної системи з можливістю контролю стану її складових, а предметом дослідження є коди з роздільним представленням складових.*

*У статті запропоновано використати методи кластерного аналізу, що, у прикладному сенсі, визначають номер кластера, до якого належить об'єкт управління. Розглянуті особливості представлення стану системи та її складових вказують на можливість використання методів м'якої кластеризації з групуванням.*

*Запропоновано та розглянуто методику синтезу сукупної оцінки стану складної системи на прикладі системи управління швидкістю вуглевидобувного комбайна. Використання кластеризації дало змогу обґрунтувати виділення окремих кластерів у просторі стану системи та принципи групування оцінок стану складових підсистем. У висновках обґрунтовано високий рівень адаптації синтезованого коду для використання у комп'ютерних мережах і системах.*

**Ключові слова:** кластер, м'яка кластеризація, стан складної системи, образ, інтегральна оцінка.

**Постановка проблеми.** У процесі створення та проектування складних виробничих систем виникають задачі ідентифікації їхнього загального стану за значеннями параметрів підсистем. Але деякі системи характеризуються невизначеним кінцевим станом або кількома можливими його значеннями. Наприклад, при вирішенні задачі вибору швидкості вуглевидобувного комбайна [1], кінцевим значенням можуть бути кілька допустимих швидкостей. Іншою особливістю є необхідність отримання дискретних значень кінцевої цільової функції

з досить обмеженого діапазону. Крім того, показники стану складових системи управління вугледобувним комбайном також можуть бути представлені обмеженою кількістю дискретних значень [2], наведених у табл. 1.

*Таблиця 1. Дискретне представлення стану підсистем*

№	Підсистема контролю	Відповідності кодовим станам
0	Навантажений стан секції МК	0 — консоль ненавантажена 1 — нормальне навантаження 2 — перевищення навантаження
1	Потенційна енергія забою	0 — підвищена напруга масиву 1 — нормальний стан масиву 2 — небезпечна напруга масиву
2	Місце розташування комбайна	0 — наявність випередження/відставання комбайна 1 — місце розташування комбайна відповідає номеру секції кріплення 2 — невідповідність місця розташування комбайна номеру секції кріплення
3	Дані маркшейдерської і гірничо-геологічної служб	0 — небезпечна зона в площині пласта 1 — відсутність небезпечних зон 2 — небезпечна зона у кривлі забою
4	Прямолінійний рух комбайна вздовж забою	0 — наявність відхилення від прямолінійності 1 — лінія забою прямолінійна 2 — лінія забою непрямолінійна
5	Прояв конвергенції	0 — жорстка посадка 1/3 кількості секцій з причин конвергенції 1 — жорстка посадка декількох секцій кріплення через утворення тріщин у шарах безпосередньої кривлі 2 — жорстка посадка всіх секцій кріплення через порушення «фактора часу»

**Аналіз останніх досліджень і публікацій.** Загальноприйнятим підходом для визначення керуючого впливу є побудова інтегральної оцінки стану системи [3], яка визначає відхилення від її «щільового» положення у n-мірному просторі стану. Особливістю формування такої інтегральної оцінки є використання систем рівнянь та отримання кінцевого результату у форматі з плаваючою комою.

Вказані вище особливості обумовлюють недоцільність обчислення інтегральної оцінки стану системи у наведеному її розумінні. Окрім цього, на вибір швидкості впливає те, яка підсистема надає вхідний вплив. При отриманні інтегральної оцінки неможливо визначити цей показник, що призводить до необхідності аналізу декількох вхідних параметрів.

Інший підхід, який дає змогу визначити керуючий вплив за множиною вхідних значень, базується на системах розпізнавання образів. Так, наприклад, задача визначення швидкості вигледобувного комбайна за допомогою нейроподібних мереж розглянута у [4].

Однак загальним недоліком використання систем розпізнавання образів є їхня досить велика складність та час вирішення задачі, що є проблематичним

у задачах реального часу, до яких відноситься більшість систем, де потрібні такі рішення. Крім того, для отримання похідної інформації потрібно кілька каналів [5], що ускладнює технічну реалізацію системи. Ці положення вказують на необхідність створення таких методів представлення оцінки стану складної системи у вигляді дискретних значень, які дають змогу визначити стан складових компонентів.

**Мета статті:** розробка методів кодування сукупної оцінки стану складної системи адаптованої до машинного аналізу з можливістю контролю стану її складових.

**Матеріали і методи.** Об'єктом дослідження є процеси кодування й аналізу стану складної системи, а предметом дослідження — коди з розділним представленням складових компонентів. Для вирішення задачі пропонується застосувати математичні підходи та методи кластерного аналізу.

**Викладення основних результатів дослідження.** У загальному представленні, сукупна оцінка стану складної системи визначається як функціонал  $F(X)$ , де  $X = \{x_1, x_2, \dots, x_N\}$ . Кожен показник  $x_n$  визначає стан відповідної підсистеми та може приймати дискретні (цілочисельні) значення. У наведеній вище задачі визначення швидкості вугледобувного комбайна, їхнє значення можна трактувати як «норма», «допустиме» або «недопустиме» з присвоєними чисельними еквівалентами, наприклад,  $p_0 = \langle 0 \rangle$ ,  $p_1 = \langle 1 \rangle$  та  $p_2 = \langle 2 \rangle$  відповідно.

Згідно з цим, усі параметри, які знаходяться в одному стані, приймають однакові значення із множини допустимих:

$$\forall x_n = \{p_0, p_1, p_2\}. \quad (1)$$

Якщо встановити однозначну відповідність між станом системи та станами її складових, то можливе потрапляння окремої складової у визначену область простору стану системи в цілому (рис. 1).

У загальному випадку, коли оцінка стану системи  $F(x_n)$  та показники стану її складових  $x_n$  можуть приймати до  $m$  станів, повинно виконуватись співвідношення:

$$\|x_n\| \leq \|F_i(x_n)\| = m. \quad (2)$$

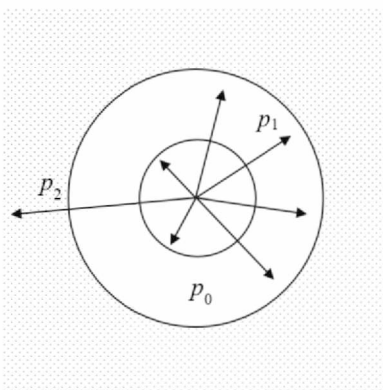


Рис. 1. Розмежування просторів стану системи

Вважатимемо, що система знаходиться у деякому  $k$ -му стані, якщо вона потрапляє в обмежену область простору значень, для якої встановлено номер  $k$ .

Введемо систему співвідношень, за якою  $F_i(x_n)$  при  $x_n = k$  набувають значення  $i=k$  та, відповідно:

$$F_i(x_n) = F_k(x_n), \text{ при } x_n = k. \quad (3)$$

Стан системи в цілому визначимо як сукупність станів її підсистем:

$$F_i(X) = \cup F_n(x_n). \quad (4)$$

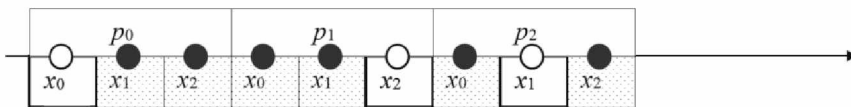
А чисельна оцінка стану системи визначається найкритичнішим значенням із параметрів її складових:

$$s = \max_n x_n. \quad (5)$$

Для представлення стану системи на основі станів підсистем, як зазначалося вище, прийнято формувати багатовимірний «образ», який можна представляти у різних формах. Для найкращого сприйняття людиною використовують декілька методів, наприклад, у вигляді пелюсткової діаграми, представленої на рис. 1, де кожній із складових виділяють окремий сектор кругової діаграми, а значення задає довжину відповідного вектора. Однак слід зазначити, що таке подання є дуже складним для машинних засобів ідентифікації, де на машинному рівні використовують найчастіше набори значень.

Для отримання машинного представлення раціонально скористатися механізмом кластеризації. За визначенням [6], «...кластер — поєднання кількох однорідних елементів, що може розглядатися самостійною одиницею, яка має визначені властивості...». Наприклад, кластери червоної, зеленої або синьої компонент (кольорові зрізи), які характеризують одну з властивостей кольорового зображення. Скориставшись цим принципом, для розглянутої задачі можна виділити кілька систем кластеризації похідної інформації — за типом підсистеми, яка її згенерувала, або за критичністю значень. У цьому разі можна виділити три кластери даних за значеннями «нормальні», «допустимі» або «недопустимі». За визначеннями, наведеними у [7], такий принцип можна віднести до кластеризації на основі групування з м'якою кластеризацією, оскільки об'єкти будуть належати цим кластерам до певної міри — доки параметри підсистеми будуть відповідати визначеним умовам.

На рис. 2 наведено модель кластерного представлення стану системи, яка характеризується множиною параметрів  $\{x_1, x_2, x_3\}$ , кожен з яких може набувати значення з множини  $\{p_0, p_1, p_2\}$ .



**Рис. 2. Геометричне представлення оцінки стану системи**

Відповідно до моделі, параметри набувають значення:  $x_0 \equiv p_0, x_1 \equiv p_2, x_2 \equiv p_1$ . Оцінка стану системи, згідно з формулою (5), дорівнює  $p_2$ .

Слід зазначити, що для визначення окремих складових та їхнього стану потрібно виокремлювати індивідуальні простори (дочірні кластери) для кожної складової. Кількість таких просторів у межах одного кластеру повинна дорівнювати кількості складових системи.

Перенесення наведеної на рис. 2 моделі на бінарну лінійну сітку дає найпростіший варіант кодування сукупної оцінки стану системи з довжиною коду 9 біт. При кодуванні простір, який відповідає елементу з відповідним станом (помічено у кластері білою точкою), замінюється на «1». Незайняті позиції (чорні точки) замінюються на «0». Відповідно, для сукупної оцінки стану системи отримуюмо код «100001010».

Для кодування параметрів, представлених у табл. 1, потрібно 18 біт (3 кластери з 6 елементами в кожному), тобто потрібно одне машинне слово довжиною 32 біти.

### **Висновки**

Запропонований процес кодування сукупного стану складної системи на основі м'якої групової кластеризації є інтуїтивно зрозумілим та простим і, відповідно, швидким у процесі як кодування, так і аналізу.

Кодування наведеної на рис. 2 моделі для бінарної сітки відбувається відповідно до співвідношення:

$$v_j = 2^{pi} \cdot 8^k, \quad (6)$$

де  $i = 0 \dots m$  — кількість станів підсистеми —  $m+1$ );  $k = 0 \dots n$  — кількість підсистем/параметрів —  $n+1$ );  $v_j$  — бінарний еквівалент положення коду стану підсистеми.

Для визначення складових, які впливають на стан системи, потрібно або послідовно накладати мазки, або скористатися зворотним механізмом від кодування за співвідношенням (6) та, використовуючи логарифмічні співвідношення, виділити відповідні компоненти.

Окремо слід відзначити гарну сумісність запропонованого методу з логічною організацією мережевого доступу у топології маркерного кільця [8], при якій кожен з абонентів (підсистема) заповнює визначене інформаційне поле мережевого повідомлення своїми даними, та використання цієї логічної організації у промислових мережах реального часу, оскільки вона забезпечує передачу даних у регламентований проміжок часу. Так, при вирішенні задачі визначення швидкості вуглевидобувного комбайна підсистема «0» заповнює «своє» поле в одному з трьох кластерів стану в повідомленні та передає його наступній підсистемі. У циклі з номером «6» повідомлення надходить до центру обробки даних, де виконується його аналіз та генерація керуючого впливу. При використанні принципу конвеєризації можна забезпечити неперервний потік актуальних даних на кожному такті, який синхронний із циклами обміну даними у комунікаційній мережі системи.

Окрім цього, запропонований метод може бути широко використаний у різноманітних системах екологічного контролю навколишнього середовища, де потрібно контролювати приналежність параметрів середовища до відповідних діапазонів, систем контролю стану ліній зв'язку або складних технічних систем [9], при реалізації сенсорних мереж систем IoT, кодуванні станів складових систем з нечіткою логікою, представленні цифрової інформації про стан ДНК у вигляді триплетів та порівнянні їх з відповідними стандартними кластерами [10].

## **Література**

1. Ларин В. Ю., Чичикало Н.И., Федоров Е. Е, Ларина Е. Ю. Савицкая Я. А. Информационная технология безопасного управления угледобывающим комплексом на основе модели нейросетевого прогнозирования. *Уголь Украины*. 2015. № 3—4. С. 63—71.
2. Савицкая Я. А. Інформаційна технологія керування вугледобувним комплексом у небезпечних зонах: дис. на здобуття наукового ступеня к.т.н.: 05.13.06 / Черкаський державний технологічний університет. Черкаси, 2018. 148 с.
3. Савицкая Я. А. Критериальный анализ параметров системы управления добычным комбайном в условиях неопределенности. *Научная мысль информационного века — 2014: десятая международ. науч.-практ. конф., Пшемьшль, 7—15 марта, 2014: тез. докл., 2014. С. 50—56.*
4. Савицкая Я. А., Чичикало Н. И. Ларин В. Ю. Методология создания нейроподобных структур системы оптимального управления угледобывающим комплексом в условиях неопределенности. *Уголь Украины*. 2014. № 9. С. 5—12.
5. Савицкая Я. А., Смолий В. В., Чичикало Н. И. Моделирование канала связи для передачи данных о состоянии угледобывающего забоя в Labview. IX Міжнародна науково-практична конференція, 03—05 жовтня 2018 р., м. Запоріжжя. С. 196—197.
6. Словари и энциклопедии. Кластер. URL: <https://dic.academic.ru/dic.nsf/ruwiki/> (дата звернення 01.11.2019).
7. Кластерний аналіз. Матеріал з Вікіпедії. URL: <https://uk.wikipedia.org/wiki/> (дата звернення 01.11.2019).
8. Офіційний сайт комітету стандартизації мережевих протоколів. Група 802. URL: <http://www.ieee802.org/5/www8025org/> (дата звернення 01.11.2019).
9. Moshenskyi A. Private Rescue Echo Beacon On SI44XX. Тези доповідей VII Міжнародної науково-технічної конференції «Проблеми Інформатизації», 13 — 15 листопада 2019 року, м. Черкаси, Том 3: секції 5—7. С. 18.
10. Попель В. А., Чумаченко С. М. Інформаційна технологія для розпізнавання образів в дослідженнях структури ДНК живих організмів. Тези доповідей VII Міжнародної науково-технічної конференції «Проблеми Інформатизації», 13 — 15 листопада 2019 року, м. Черкаси, Том 3: секції 5—7. С. 19.