

## **ЗАСТОСУВАННЯ КЛАСТЕРНОГО АНАЛІЗУ У АВТОМАТИЗОВАНИХ СИСТЕМАХ КОНТРОЛЮ ЯКОСТІ ЕЛЕКТРОЕНЕРГІЇ**

**Светелік О.О., Солодовник А.О.**

***Національний технічний університет  
«Харківський політехнічний інститут», м. Харків***

У теперішній час в Україні триває процес побудови автоматизованої системи контролю якості електричної енергії у енергосистемі України. При наявності великої кількості приладів реєстрації порушень постає питання множення зафіксованих подій, так як усі елементи енергосистеми працюють у взаємозв'язку. Таким чином, наприклад, провал напруги не може існувати лише на окремому приєднанні, а спостерігається у тому чи іншому вигляді на значній ділянці електричної мережі. Для підвищення зручності та швидкості обробки інформації про порушення доцільно групувати та об'єднувати записи з різних пристроїв про одну і ту саму подію. Найбільш ефективно це можна зробити за допомогою апарату кластерного аналізу. Кластеризація - це процес розбиття безлічі об'єктів на групи, які називаються кластерами. У середині кожної групи повинні виявитися «схожі» об'єкти, а об'єкти різних групи мають відрізнятися якомога більше. Головна відмінність кластеризації від класифікації у тому, що перелік груп чітко не заданий й визначається у процесі роботи алгоритму. Головний принцип полягає у том, що для кожної пари об'єктів розраховується «відстань», яка є мірою схожості. Існує велика кількість метрик, найпоширеніша з яких – евклідова відстань. Така метрика ефективна для врахування електричної близькості, або різниці у часі між зареєстрованими подіями. Але вона не підходить при врахуванні типу події, наприклад провалу напруги, несиметрії, або гармонік. У цьому випадку доцільно випростовувати відстань Чебишева, або ступінну відстань.

Відстань Чебишева застосовується, коли потрібно визначити два об'єкти як «різні», якщо вони відрізняються за якоюсь однією координатою, розраховується за формулою:

$$\rho(x, x') = \sum_i^n |x_i - x'_i|.$$

Ступінна відстань використовується для випадків, коли необхідно збільшити або зменшити вагу, що стосується розмірності, для якої відповідні об'єкти сильно відрізняються:

$$\rho(x, x') = \sqrt[p]{\sum_i^n (x_i - x'_i)^p}.$$

Використання наведених метрик дозволяє покращити якість побудови кластерів подій порушень якості електричної енергії у автоматизованих системах моніторингу.