

УДК 621.311.25:519.816

П.Ф. Буданов, К.Ю. Бровко

Украинская инженерно-педагогическая академия, Харьков

## МЕТОД КЛАСТЕРНОГО АНАЛИЗА ДЛЯ ОБРАБОТКИ ИНФОРМАЦИОННОГО ПРОСТРАНСТВА В АВТОМАТИЗИРОВАННЫХ ТРЕНАЖЕРАХ ПО ПОДГОТОВКЕ ОПЕРАТИВНОГО ПЕРСОНАЛА АЭС

Рассмотрены существующие методы кластерного анализа и показано, что для задачи обработки информации об изменениях и отклонениях параметров технологического процесса более целесообразно использовать методы иерархического дивизионного и алгоритмического типа, которые позволяют объединять и связывать технологические параметры в кластер-кластерные агрегации. Для описания процесса обработки объема информации в автоматизированных тренажерах по подготовке оперативного персонала АЭС предложена математическая модель изменения параметров технологического процесса в степенной зависимости от количественной величины фрактальной размерности, характеризующей степень заполнения информационного пространства, а также предложен подход к разработке алгоритма принятия решений оперативным персоналом в нештатных аварийных ситуациях на АЭС, для выбора оптимальных вариантов решения.

**Ключевые слова:** автоматизированные системы обучения, оперативный персонал, штатные и нештатные ситуации, модель и алгоритм тренажера, кластерный анализ.

### Введение

#### Постановка проблемы и анализ публикаций.

Существующие сегодня методы обработки информации об аварийных ситуациях на потенциально опасных объектах (ТЭС и АЭС), требуют дальнейшего усовершенствования по эффективности и надежности выявления опасных нештатных процессов, происходящих в реальных масштабах пространства и времени [1, 2]. Рассмотрение методов обработки сигналов, таких как метод «Дерево событий»; «Дерево процессов»; «Модель в сигнатуре» показало, что они в недостаточной мере, показывают реальную картину происходящих событий.

Увеличение числа контролируемых и регулируемых технологических параметров АЭС обусловило увеличение объема различной информации, которую должен принять и переработать оперативный персонал [1].

В информационном пространстве выполняются следующие функции: оперативный контроль о ходе технологического процесса, технологическая сигнализация, расчет технико-экономических показателей, определение достоверности информации, диагностика состояния оборудования, регистрация аварийных ситуаций [1, 2].

В современных теориях принято рассматривать информационное пространство, как стохастическое, т.е. имеющее свойство неопределенности и случайности, поэтому современные методы обработки информации допускают недостоверность и неточность в обработке.

Для принятия решения (ПР) оперативному персоналу АЭС необходимо адекватно оценивать поступающие параметры технологического процесса

АЭС. Во многих моделях в информационном пространстве изучаются структурные связи между параметрами, несущими информацию о технологическом процессе (ТП), но они в недостаточной мере описывают поступающую информацию для оптимального выбора варианта принятия решения [2].

Одной из наиболее характерных особенностей управления энергоблоками АЭС является чрезвычайно большой объем информации [1], которую необходимо получить, переработать и реализовать при управлении электрооборудованием и для обеспечения его надежной и экономичной работы в различных режимах (табл. 1).

Таблица 1  
Объемы контроля и управления АЭС

Мощность энергоблока, МВт	Количество единиц		
	Контролируемые параметры	Объекты диспетчерского управления	Регулируемые параметры
1000	4500	1500	900

Режимы эксплуатации АЭС можно разделить на следующие группы.

Рассмотрим основные из них:

- режим нормальной эксплуатации: плановые пуски и остановки энергоблоков; стационарная работа на разных уровнях мощности; переходные режимы; расхолаживание; перегрузка ядерного топлива;
- аварийный режим эксплуатации, который может привести к повреждению ТВЭЛов и загрязнению помещений АЭС (окружающей среды).

Особенно трудные условия складываются во время нестационарных условий работы (н-р, пуск или остановка блока), а также во время аварийных ситуаций, когда оперативному персоналу необходи-

мо манипулировать десятками регулирующих и запорных органов при одновременном контроле за многочисленными технологическими параметрами.

На основе энергетических характеристик и текущей информации о состоянии оборудования и параметрах технологического процесса решаются такие задачи, как: оптимальное распределение активной нагрузки между энергоблоками; выбор оптимального состава действующего оборудования; распределение видов и количества топлива и т.д.

Чрезвычайно большой объем информации, который необходимо собрать, переработать, хранить и реализовать при управлении современным энергоблоком АЭС, обусловил появление принципиально новых методов сбора, обработки и представления информации с использованием качественно новых технических средств и технологического оборудования.

Сегодня существует проблема по организации потоков информации к оперативному персоналу и командной информации от него, решение которой обеспечило бы оптимальные действия оператора по принятию решений в нештатных аварийных ситуациях [1, 2].

Как известно [1 – 4], информационные модели описывают информационные связи, в том числе и количественные характеристики потоков информации между уровнями и в пределах уровневой структуры. На основе информационной модели определяются необходимые характеристики устройств ввода-вывода информации; производятся оценки памяти для хранения информации; определяется периодичность решения функциональных задач.

Вся поступающая информация к оператору доносится в виде информационных сигналов о состоянии параметров технологического процесса.

Применение математических методов моделирования применительно к информационным каналам обусловлено необходимостью диагностики искажений сигнала и его коррекции с целью повышения информационной надежности.

Основная проблема диагностики сигналов связана с развитием методов выявления искажений и оценки их параметров, что невозможно без предварительного построения математической модели информационного сигнала.

Среди известных методов анализа информационного пространства, авторами предложено использовать методы кластерного анализа [3, 5, 6].

Как известно [3, 5, 6], при моделировании объема информационного пространства все чаще используется синергетический фрактальный подход, одним из основных свойств которого является свойство самоподобия, упорядоченности.

Синергетика [5] и теория фракталов [6] тесно связана с кластерным анализом, который решает

задачу выделения компактных групп объектов с близкими свойствами и характеристиками, в данном случае о состоянии параметров технологического процесса.

**Целью статьи** является разработка математической модели процесса принятия решений оперативным персоналом АЭС в нештатных аварийных ситуациях на основе аппарата теории фракталов и синергетики с использованием методов кластерной обработки информационного пространства.

## Основной материал

Математическая модель информационного пространства представляет собой множество фрактальных кластерных структур.

Методы кластерного анализа можно разделить на две группы: иерархические; неиерархические.

Каждая из групп включает множество подходов и алгоритмов. Используя различные методы кластерного анализа, таким образом, получают различные варианты принятия решений для одних и тех же параметров технологического процесса, в том числе тех, которые приводят к нештатным аварийным ситуациям на АЭС.

**Методы объединения или связи.** Когда каждый объект представляет собой отдельный кластер, расстояния между этими объектами определяются выбранной мерой. Для определения расстояния между кластерами существуют различные правила, называемые методами объединения или связи для двух кластеров.

**Метод ближнего соседа или одиночная связь.** Здесь расстояние между двумя кластерами определяется расстоянием между двумя наиболее близкими объектами (ближайшими соседями) в различных кластерах. Этот метод позволяет выделять кластеры сколь угодно сложной формы при условии, что различные части таких кластеров соединены цепочками близких друг к другу элементов. В результате работы этого метода кластеры представляются длинными «цепочками» или кластерами, «сцепленными вместе» только отдельными элементами (отдельные характеристики параметров), которые случайно оказались ближе остальных друг к другу.

**Метод наиболее удаленных соседей** или полная связь. Здесь расстояния между кластерами определяются наибольшим расстоянием между любыми двумя объектами в различных кластерах (т.е. «наиболее удаленными соседями»). Метод хорошо использовать, когда объекты действительно происходят из различных «кластерных групп». Если же кластеры имеют в некотором роде удлиненную форму или их естественный тип является «цепочечным», то этот метод не следует использовать.

**Метод Варда.** В качестве расстояния между кластерами берется прирост суммы квадратов рас-

стояний объектов до центров кластеров, получаемый в результате их объединения. В отличие от других методов кластерного анализа для оценки расстояний между кластерами, используются методы дисперсионного анализа. На каждом шаге алгоритма объединяются такие два кластера, которые приводят к минимальному увеличению целевой функции, т.е. внутригрупповой суммы квадратов. Этот метод направлен на объединение близко расположенных кластеров и «стремится» создавать кластеры малого размера.

Метод невзвешенного попарного среднего (метод невзвешенного попарного арифметического среднего). В качестве расстояния между двумя кластерами берется среднее расстояние между всеми парами объектов в них. Этот метод следует использовать, если объекты действительно происходят из различных «кластерных групп», в случаях присутствия кластеров «цепочного» типа, при предположении неравных размеров кластеров.

Метод взвешенного попарного среднего (метод взвешенного попарного арифметического среднего). Этот метод похож на метод невзвешенного попарного среднего, разница состоит лишь в том, что здесь в качестве весового коэффициента используется размер кластера (число объектов, содержащихся в кластере). Этот метод рекомендуется использовать именно при наличии предположения о кластерах разных размеров.

Невзвешенный центроидный метод (метод невзвешенного попарного центроидного усреднения). В качестве расстояния между двумя кластерами в этом методе берется расстояние между их центрами тяжести.

Взвешенный центроидный метод (метод взвешенного попарного центроидного усреднения). Этот метод похож на предыдущий, разница состоит в том, что для учета разницы между размерами кластеров (числе объектов в них), используются веса. Этот метод предпочтительно использовать в случаях, если имеются предположения относительно существенных отличий в размерах кластеров.

Применение иерархических методов кластерного анализа. Суть применения иерархической кластеризации состоит в последовательном объединении меньших кластеров в большие или разделении больших кластеров на меньшие (рис. 1), что приводит соответствующим уменьшением числа кластеров.

В начале работы алгоритма все объекты являются отдельными кластерами. На первом шаге наи-

более похожие объекты объединяются в кластер. На последующих шагах объединение продолжается до тех пор, пока все объекты не будут составлять один кластер.

#### Иерархические дивизимные (делимые) методы.

Эти методы являются логической противоположностью агломеративным методам. В начале работы алгоритма все объекты принадлежат одному кластеру, который на последующих шагах делится на меньшие кластеры, в результате образуется последовательность расщепляющих групп (рис. 1).

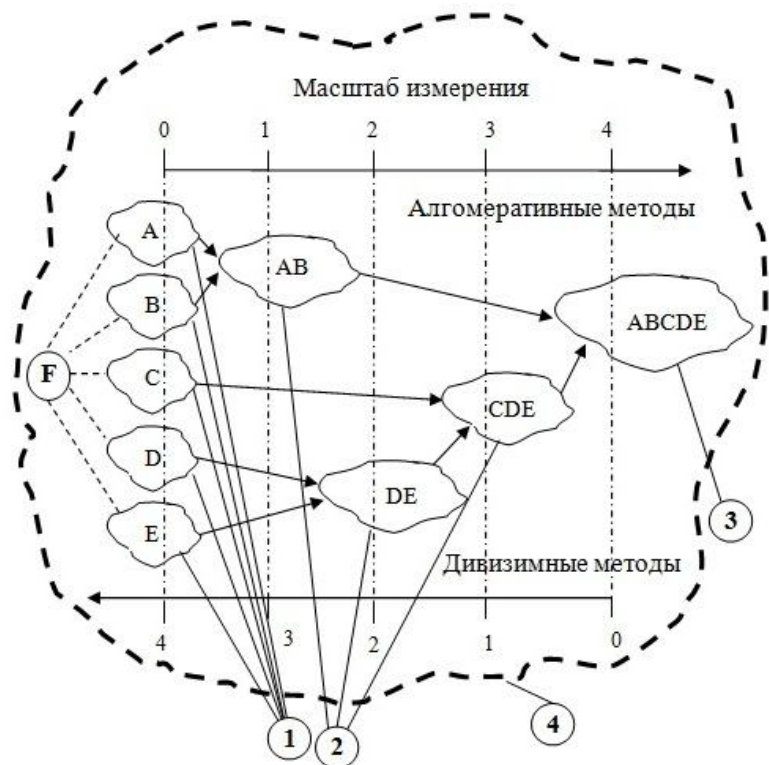


Рис. 1. Иерархические методы кластерного анализа  
1 – параметры технологического процесса;  
2 – кластерные группы;  
3 – кластер-кластерные агрегации;  
4 – объем информационного пространства

Авторами показано, что применение иерархических и неиерархических методов кластерного анализа применительно к задаче анализа объема информационного пространства (ОИП), которое заполнено информационными параметрами о состоянии технологического процесса, является предпочтительным, так как ОИП рассматривается как в пространстве, так и во времени.

В ходе технологического процесса происходит изменение различных характеристик технологических параметров, которые можно представить в виде отдельных кластеров (A, B, C, D, E).

Далее образованные кластеры технологических параметров взаимодействуют между собой создают кластерные группы по определенным признакам и

характеристикам параметров, и таким образом образуют кластер-кластерные агрегации, тем самым, заполняя объем информационного пространства с признаками аварийности.

На основании рассмотренного метода кластерного анализа по объединению технологических параметров в кластер-кластерную агрегацию, авторами было предложено для моделирования алгоритма по принятию решения оперативным персоналом в нештатных аварийных ситуациях применять сетевое моделирование, как показано на рис. 2.

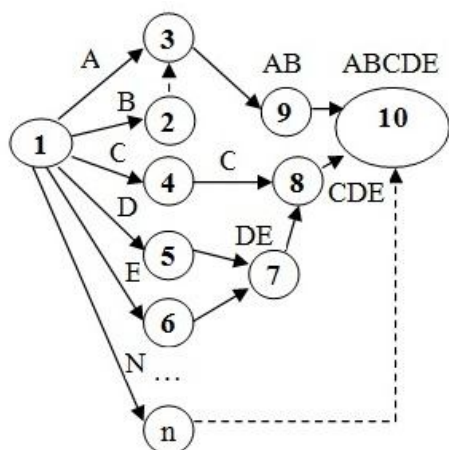


Рис. 2. Модель алгоритма принятия решений оперативным персоналом в нештатных ситуациях

Из рис. 2 видно, что в исходном состоянии все технологические параметры имеют нормированные характеристики, которые в последующем могут изменяться, что приводит к формированию кластерных групп или отдельных кластеров (н-р: кластерные группы – AB и CDE), которые и формируют кластер-кластерную агрегацию (н-р: ABCDE). Таким образом, в таком алгоритме, как видно из рис. 2, представлены пять вариантов принятия решений по данной информации (н-р: варианты – 1,2,3,10; 1,2,10;)

Кроме того, кластер-кластерная агрегация имеет следующие математические характеристики, которые позволяют определить фрактальную размерность  $d_f$ : центр кластера (это среднее геометрическое место точек в пространстве переменных); радиус кластера (максимальное расстояние точек от центра кластера), среднее квадратическое отклонение, размер кластера.

Кластеры также могут быть перекрывающимися. Такая ситуация возникает, когда обнаруживается перекрытие кластеров. В этом случае невозможно при помощи математических процедур однозначно отнести объект к одному из двух кластеров. Такие объекты называют спорными. Спорный объект – это объект, который по мере сходства может быть отнесен к нескольким кластерам.

Размер кластера может быть определен либо по радиусу кластера, либо по среднее квадратическому

отклонению объектов для этого кластера. Объект относится к кластеру, если расстояние от объекта до центра кластера меньше радиуса кластера. Если это условие выполняется для двух и более кластеров, объект является спорным.

Работа кластерного анализа опирается на два предположения.

Первое предположение – рассматриваемые признаки объекта в принципе допускают желательное разбиение (совокупности) объектов (пространства) на кластеры.

Второе предположение – правильность выбора масштаба или единиц измерения признаков.

Классический подход к диагностике состоит в анализе информационного сигнала на приемной стороне и определении частоты появления ошибочных символов и коэффициента ошибок модуляции. Однако указанные параметры дают представление лишь о степени ухудшения качества сигнала, но не о возможных причинах его снижения, более того, для получения достоверной оценки необходимо длительное время наблюдения, т.е. ухудшение качества информационного сигнала регистрируется с большой задержкой.

Метод прогнозирования последствий и модель автоматизированной системы поддержки действий оперативного персонала при возникновении аварийных ситуаций базируются на рассмотрении сложной системы обработки информационного пространства, включающей в себя технологические параметры, оперативно – диспетчерскую службу, аппаратную систему контроля технического состояния оборудования как единую информационную структуру.

Необходимо осуществить синтез оптимальных решений, обеспечивающих управление имеющимися в распоряжении оперативного персонала, принимающего решения, ресурсами в целях, в случае возникновения аварии, локализации ее масштабов, минимизации потерь располагаемых ресурсов.

Концепция по разработке математической модели в процессе принятия решений оперативным персоналом АЭС при возникновении и развитии аварий основана на последовательности преобразований хранящейся и поступающей информации.

Эту модель предложено представить в виде выражения (1), как функции ряда упорядоченных элементов (признаков нештатных сценариев) A, B, C, D, E:

$$MM_{кл} = F(A, B, C, D, E \dots N)^{df}, \quad (1)$$

где A – параметры о состоянии средств технического обеспечения информации; B – параметры, определяющие отклонение в технологическом процессе, которые приводят к нештатной ситуации в пространстве и времени; C – параметры, характеризующие масштабы и величины параметров аварийного процесса; D – параметры, определяющие тип нештатного аварийного события;  $E = [E_{оп}, E_{спш}]^t$  –

величина параметров, характеризующих действия оперативного персонала ( $E_{оп}$ ) в различных нештатных ситуациях и список параметров ( $E_{спнш}$ ), формируемый динамически на основе данных выбора оптимального варианта решения.

Предположим, что рассматриваемая система обработки информационного пространства (СОИП) состоит из  $k$  подсистем, тогда для любой  $k$ -й подсистемы процесс принятия решений оперативным персоналом определяется упорядоченными элементами математической модели  $MM_k$ , что является основанием для применения аппарата фрактальной теории [6] и методов кластерного анализа для группировки вышеуказанных элементов (признаков) в соответствующие кластеры, т.е. образования кластер-кластерной агрегации, что позволяет по новому представить процесс обработки информации и представить ее в степенной зависимости от количественной величины – фрактальной размерности  $d_f$  [6].

Необходимо также отметить, что основной количественной величиной, которая характеризует степень незаполненности информационного пространства, является фрактальная размерность.

Также впервые предложено рассматривать систему СОИП, как открытую синергетическую систему [3, 5], при которой совокупность (сумма) информационных элементов превышает объем информационного пространства, как показано на рис. 3.

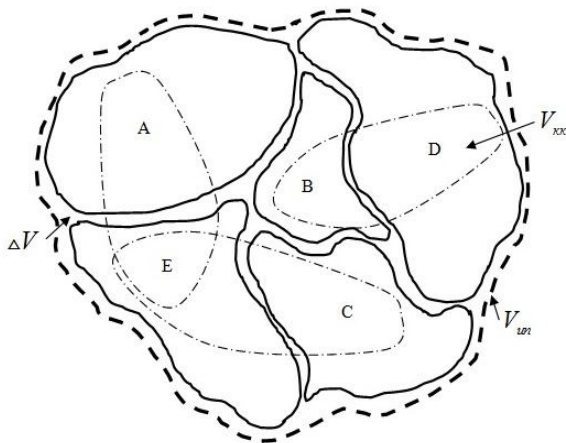


Рис. 3. Графическое представление формирования кластер-кластерной агрегации в объеме информационного пространства

Из рис. 3 видно, кластеры отдельных параметров образуют определенный объем кластер-кластерных агрегаций ( $V_{кк}$ ), откуда следует, что полный объем информационного пространства ( $V_{инп}$ ) имеет незаполненные частицы пространства ( $\Delta V$ ), т.е.  $\Delta V = V_{инп} - V_{кк}$ , а, следовательно, коэффициент заполнения информационного пространства  $k_{зи} = V_{кк} / V_{инп}$ .

Для количественной характеристики объема информационного пространства на основе кластер-

ного анализа авторами предложено использовать аппарат теории фракталов.

Поэтому количественной физической величиной, которая характеризует степень заполнения всего объема информационного пространства ( $V_{инп}$ ) соответствующими параметрами (A, B, C, D, E) в соответствии с теорией фракталов, является фрактальной размерность ( $d_f$ ) кластер-кластерной агрегации.

Таким образом, для описания процесса обработки информации и принятия решения, есть основание в применении кластерного анализа и аппарата фрактальной теории, т.е. достижение связи изменения фрактальной размерности ( $d_f$ ) от динамики изменения характеристик хотя бы одного параметра (A, B, C, D, E), которые могут привести к нештатным аварийным ситуациям, т.е. определяется выражением (2):

$$V_{кк} = (V_A + V_B + V_C + V_D + V_E)^{d_f} \quad (2)$$

или 
$$V_{кк} = (\beta V_{гр})^{d_f}, \quad (3)$$

где  $V_{кк}$  – объем кластер-кластерной агрегации (общее число параметров в информационном потоке);  $V_{гр}$  – объем групп объединенных кластеров (параметров);  $d_f$  – фрактальная размерность информационного массива;  $\beta$  – масштаб измерения.

Из выражений (2), (3) следует, что соотношение между количеством параметров в объеме кластер-кластерной агрегации  $V_{кк}$ , несущих информацию о нештатном технологическом процессе, и кластеров, входящих в группы кластеров  $V_{гр}$ , проявляется при сохранении фрактальных свойств внутренней структуры информационного пространства при изменении масштабов его внешнего рассмотрения.

Теория фракталов [6] широко применяется как подход к исследованию рядов наблюдений, позволяющий получать важные характеристики параметров технологических процессов, не вдаваясь в подробный анализ их внутренней структуры.

## Выводы

1. Проанализированы современные методы кластерного анализа по обработке информационного пространства и выделены наиболее эффективные методы объединения (связи) технологических параметров в кластер-кластерные агрегации.

2. Используя методы кластерного анализа, на основе сетевого моделирования разработан алгоритм принятия решений оперативным персоналом в нештатных аварийных ситуациях, который позволяет выбрать оптимальный вариант решения.

3. Предложена математическая модель процесса принятия решений оперативным персоналом АЭС при возникновении и развитии аварий основе данных кластерного анализа.

## Список літератури

1. Артюх С.Ф. Основы автоматизированных систем управления энергогенерирующими установками электростанций / С.Ф. Артюх, М.А. Дуэль, И.Г. Шелепов. – Х., 1998. – 332 с.
2. Буданов П.Ф. Анализ современного состояния и перспективы развития автоматизированных систем по подготовке оперативного персонала АЭС / П.Ф. Буданов, К.Ю. Бровко // Системы обработки информации. – Х.: ХУПС, 2013. – Вып. 9(107). – С. 263-269.
3. Буданов П.Ф. Синергетический подход к разработке модели принятия решения оперативным персоналом АЭС в нестандартных ситуациях. / П.Ф. Буданов, К.Ю. Бровко // Системы обработки информации. – Х.: ХУПС, 2013. – Вып. 1(108). – С. 256-262.
4. Ландэ Д.В. Объектно-статистический анализ информационных потоков / Д.В. Ландэ, В.Н. Фурашев // Открытые информационные и компьютерные технологии. – Х.: НАКУ «ХАИ», 2007. – Вып. 35. – С. 133-137.
5. Хакен Г. Синергетика / Г. Хакен. – М.: Мир, 1980. – 404 с.
6. Федер Е. Фракталы / Е. Федер. – М.: Мир, 1991. – 260 с.

Поступила в редколлегию 11.02.2013

Рецензент: д-р техн. наук Г.И. Канюк, Украинская инженерно-педагогическая академия, Харьков.

### МЕТОД КЛАСТЕРНОГО АНАЛІЗУ ДЛЯ ОБРОБКИ ІНФОРМАЦІЙНОГО ПРОСТОРУ В АВТОМАТИЗОВАНИХ ТРЕНАЖЕРАХ ПО ПІДГОТОВЦІ ОПЕРАТИВНОГО ПЕРСОНАЛУ АЕС

П.Ф. Буданов, К.Ю. Бровко

Розглянуті існуючі методи кластерного аналізу і показано, що для завдання обробки інформації про зміни і відхилення параметрів технологічного процесу, доцільніше необхідно використовувати методи ієрархічного дивізіонного і алгомеративного типу, які дозволяють об'єднувати і зв'язувати технологічні параметри в кластер-кластерні агрегації. Для опису процесу обробки об'єму інформації в автоматизованих тренажерах з підготовки оперативного персоналу АЕС, запропонована математична модель зміни параметрів технологічного процесу в статичній залежності від кількісної величини фрактальної розмірності, що характеризує міру заповнення інформаційного простору, а також запропонований підхід до розробки алгоритму прийняття рішень оперативним персоналом в нестандартних аварійних ситуаціях на АЕС, для вибору оптимальних варіантів рішення

**Ключові слова:** автоматизовані системи навчання, оперативний персонал, штатні і нестандартні ситуації, модель і алгоритм тренажера, кластерний аналіз.

### METHOD OF CLUSTER ANALYSIS FOR TREATMENT OF INFORMATIVE SPACE IN THE AUTOMATED TRAINERS ON PREPARATION OF OPERATIVE PERSONNEL OF NUCLEAR POWER PLANT

P.F. Budanov, K.Yu. Brovko

The existent methods of cluster analysis are considered it is shown that for the task of treatment of information about changes and rejections of parameters of technological process, it is more expediently necessary to use the methods of hierarchical devising and alglomerativ type, which allow to unite and link technological parameters in the rlastrer-rlaster aggregating. For description of process of treatment of volume of information in the automated trainers on to preparation of operative personnel of nuclear power plant, the mathematical model of change of parameters of technological process is offered in sedate dependence on the quantitative size of fractal dimension, characterizing the degree of filling of informative space, and also offered approach to development of algorithm making decision an operative personnel in nonpermanent emergency situations on nuclear power plant, for the choice of optimal variants of decision

**Keywords:** automated departmental teaching, operative personnel, regular and nonpermanent situations, model and algorithm of trainer, cluster analysis.