

СООБЩЕНИЯ
ОБЪЕДИНЕННОГО
ИНСТИТУТА
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ
ДУБНА

13-85-795

Ю. Н. Пепёлышев

ИССЛЕДОВАНИЕ
АЛГОРИТМА РАСПОЗНАВАНИЯ ОБРАЗОВ
В ЗАДАЧЕ КЛАССИФИКАЦИИ СПЕКТРОВ
ШУМОВ РЕАКТОРА ИБР-2

1985

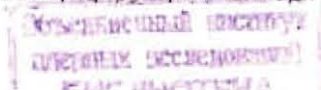
ВВЕДЕНИЕ

Одно из направлений в диагностике работы ядерного реактора связано с анализом шумовых характеристик различных реакторных параметров^{1,2}. В общем случае по характеру шумов некоторого числа реакторных параметров требуется определить состояние реактора /"плохо-хорошо"/ и охарактеризовать это состояние набором наиболее важных для диагностики реактора параметров /например, уровнем средней мощности, флуктуациями мощности и т.д./. Такой подход требует предварительной классификации реакторных шумов, т.е. выделения типовых классов спектральных структур в колебаниях мощности /и /или/ в шумах других реакторных параметров/ и установления адекватной связи выделения классов с определенным состоянием реактора. Число и тип параметров, идентифицируемых по шумам мощности, во многом зависит от особенностей реактора и реакторных шумов, от метода получения спектральных оценок, а также от выбора параметров классифицирующего алгоритма. Кроме того, задача классификации, как правило, связана с обработкой большого объема исходной информации. Например, для реактора ИБР-2 требуемый для обработки объем исходной экспериментальной информации составляет $\sim 10^6$ и более слов, а информация о спектрах только одного реакторного параметра - $\sim 3 \cdot 10^4$ слов. Вместе с тем число реакторных параметров, непосредственно влияющих на установившуюся мощность реактора, достаточно велико. Отсюда ясно, что для классификации шумов реактора ИБР-2 необходимо, во-первых, сжатие обрабатываемой информации, и, во-вторых, проведение предварительного обучения для определения таких параметров алгоритма, при которых ошибка классификации была бы минимальна.

В данной работе исследуется один из вариантов полного алгоритма классификации, включающий способ получения спектральных оценок, формализованный алгоритм сжатия информации, а также выбор параметров классифицирующего алгоритма.

1. ИСХОДНЫЕ ДАННЫЕ И ВЫБОР ИНФОРМАТИВНЫХ ПРИЗНАКОВ СПЕКТРАЛЬНЫХ ОЦЕНОК

Возможность оптимизации полного алгоритма классификации решалась на модельной задаче с использованием процедуры обучения для флуктуаций мощности, измеренных при известных состояниях реактора. В основе классифицирующего алгоритма был принят эври-



стический алгоритм распознавания образов ISODATA². Классификация шумов осуществлялась по их спектрам путем выделения геометрически интерпретируемых инвариантных признаков, отражающих сохраняющиеся особенности в форме отдельных групп спектров /кластеров/. В качестве классификатора использовалась мера евклидова расстояния. Ошибка классификации зависела в этом случае от параметров ISODATA и от погрешностей спектральных оценок. Исходная экспериментальная информация представляла собой массив данных $\{\theta_{\ell}\}_j$, где $\ell = 1, 8192$, $j = 1, 19 / 8192$ - число последовательных значений энергии импульсов мощности, измеренных при j -ом состоянии реактора/.

Измерения $\{\theta_{\ell}\}_j$ проводились при частоте повторения импульсов мощности 5 с^{-1} , т.е. максимальная частота полученных при обработке $\{\theta_{\ell}\}_j$ спектров \bar{s}_j , составляла 2,5 Гц. Условия измерений отличались уровнем средней мощности W и суммарным расходом теплоносителя через активную зону G . Значения W и G для каждого i -го состояния реактора представлены в табл.1.

Таблица 1
Значения средней мощности реактора W , суммарного расхода натриевого теплоносителя через активную зону G и относительное стандартное отклонение σ_W/W флуктуаций мощности для j -го состояния реактора

i	W , МВт	G , м ³ /ч	σ_W/W
1	0,4	40	1,93
2	0,4	40	1,87
3	0,4	40	1,86
4	0,4	66	1,84
5	0,4	70	1,67
6	0,4	70	1,71
7	0,4	70	1,67
8	0,16	26	1,94
9	0,24	26	2,13
10	0,24	26	1,99
11	1,0	80	2,24
12	1,0	80	2,09
13	1,0	80	2,19
14	1,0	80	1,6
15	1,0	80	2,96
16	1,0	80	3,11
17	1,0	80	3,09
18	1,0	80	3,88
19	1,0	80	3,12

Для классификации использовались оценки нормированных спектральных плотностей $S_i(f_i) = S_i(f_i) / D_{W_i}$, $i = 1, N$, N - полное число точек спектральной оценки, D_{W_i} - дисперсия j -ой оценки. Таким образом, объем классифицируемой выборки составлял $19 \times N$. Предварительное изучение структуры спектров, полученных с использованием быстрого преобразования Фурье, показало, что значение N следует выбрать равным 128 или 256. Большее число /т.е. 512/ не добавляло новой информации, а меньшее /т.е. 64/ приводило к существенно неравномерному по f_i отрицательному смещению спектральных оценок и, соответственно, - к общему искажению их формы. В зависимости от метода получения S_i /при фиксированном значении N / спектральные оценки отличаются двумя основными параметрами: разрешающей способностью Δf и статистической

погрешностью $\epsilon = \sigma_{S_i} / S_i$. Для оценки допустимых искажений формы спектров при классификации последних по алгоритму ISODATA исследовалось несколько вариантов $S_i(f_i)$, отличающихся значениями Δf и ϵ /см. табл.2/. Выбранные значения Δf и ϵ зависели от принятого ранее оптимального значения N и от способа обработки исходной информации. Спектральные оценки были получены по методу Кули-Тьюки с использованием или временного окна данных Кайзера-Бесселя³, или прямоугольного временного окна. На рис.1,2 показаны две наиболее отличающиеся спектральные выборки /при $\epsilon = 25\%$ и $\epsilon = 9\%$ /.

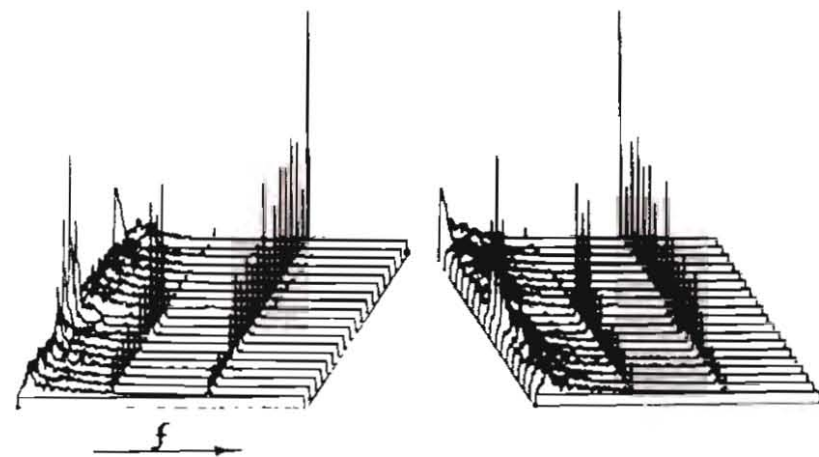


Рис.1. Нормализованная спектральная плотность колебаний мощности $N = 257$, $\epsilon = 25\%$, $\Delta f = 0,012$ Гц

Для дополнительного сжатия информации применялась формализованная процедура выбора наиболее информативных признаков спек-

Распределение спектров по кластерам для некоторых значений относительного стандартного отклонения ϵ и разрешающей способности Δf спектральных оценок

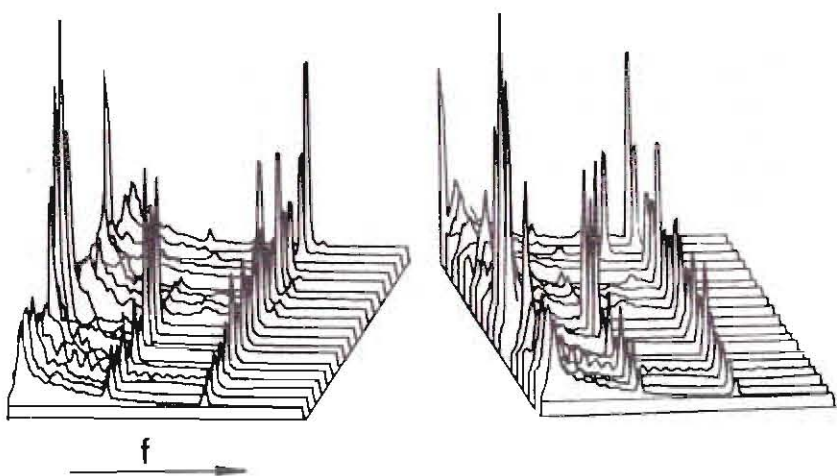


Рис. 2 Нормализованная спектральная плотность колебаний мощности. $N = 129$, $\epsilon = 9\%$, $\Delta f = 0,05$ Гц.

тров. В качестве меры информативности признака, за который была принята спектральная плотность на частоте f_i , принималась энтропия распределения значений признака $H(f_i)$, $i = \overline{1, N}$:

$$H(f_i) = - \sum_{k=1}^m P_k \log P_k \quad /1/$$

где P_k - вероятность нахождения значения признака на частоте f_i в k -ом интервале его изменения, m - число интервалов разбиения значений признака. На рис. 3, 4 для примера показано изменение энтропии $H(f_i)$ и дисперсии признаков $D(f_i)$ для спектральной выборки, соответствующей рис. 1 /см. пункт 1 табл. 2/. В качестве информативных признаков принимались значения спектральной плотности на частоте f_n , для которой одновременно удовлетворялись следующие условия:

$$H(f_n) \leq \bar{H} - d_1 \sigma_H,$$

$$D(f_n) > \bar{D} + d_2 \sigma_D.$$

где \bar{H}, σ_H - среднее значение и, соответственно, - стандартное отклонение распределения энтропии; \bar{D}, σ_D - среднее значение и, соответственно, - стандартное отклонение распределения дисперсии признаков, d_1, d_2 - коэффициенты, учитывающие форму распределения H и D , значение которых для всех вариантов $S_i(f_i)$, указанных в табл. 2, составляло 0,25. В зависимости от вида спек-

№	Объем спектральной выборки		ϵ , %	Δf^* , Гц		Распределение спектров по кластерам
	Исходная выборка	После отбора информации				
1	19 x 257	19 x 15	25	0,012	-	1 - 6,7 2 - 8,9,10 3 - 18 4 - 1,2,3,4 5 - 11,13,14,15,16,17 6 - 19 7 - 5,12
2	19 x 129	19 x 10	18	0,024	-	1 - 6,7 2 - 8,9,10 3 - 18 4 - 1,2,3,4,5 5 - 11,12,13,14,15,16,17,19
3	19 x 257	19 x 17	13	0,025	+	1 - 6 2 - 8,9,10 3 - 18 4 - 1,2,3,4,5 5 - 11,12,13,14,15,16,17 6 - 19
4	19 x 129	19 x 11	9	0,05	+	1 - 6 2 - 8,9,10 3 - 18 4 - 1,2,3,4,5,7 5 - 11,12,13,15,16,17 6 - 14,19

Примечание: знак + указывает на использование в данном варианте обработки временного окна данных Кайзера-Бесселя

* Δf - взята по уровню 6Db спектрального окна.

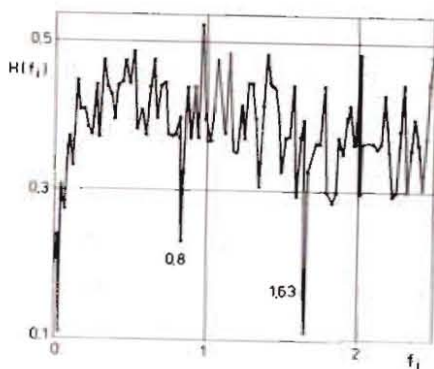


Рис.3. Зависимость энтропии распределения признака $H(f_i)$ от частоты f_i . $N = 257$, $\epsilon = 25\%$, $\Delta f = 0,012$ Гц.

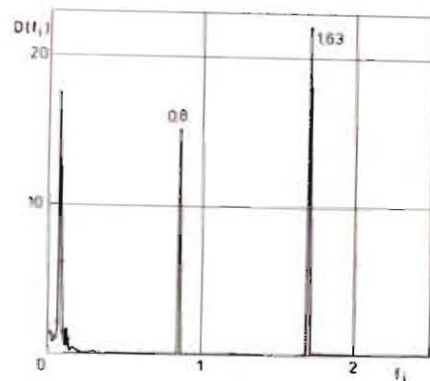


Рис.4. Зависимость дисперсии признака $D(f_i)$ от частоты f_i . $N = 257$, $\epsilon = 25\%$, $\Delta f = 0,012$ Гц. Цифрами на рисунке отмечены значения характерных частот.

тральной выборки число информативных признаков спектров менялось от 11 до 17 /см.табл.2/, т.е. более чем на порядок оказалось меньше числа исходных точек спектров.

3. ВЫБОР ОПТИМАЛЬНЫХ ПАРАМЕТРОВ КЛАССИФИЦИРУЮЩЕГО АЛГОРИТМА И АНАЛИЗ РЕЗУЛЬТАТОВ ОБРАБОТКИ

В процессе работы алгоритма ISODATA определялись: число кластеров /групп спектров/, номера спектров, принадлежащих данному кластеру, а также форма спектра, соответствующая центру каждого кластера. Для всех наборов $S_i(f_i)$, полученных с разными значениями ϵ и Δf , оптимизировались следующие параметры алгоритма: среднеквадратичное отклонение спектров в кластере θ_s , параметр θ_c , характеризующий компактность кластеров, а также желаемое число кластеров N_c . Оптимизация параметров осуществлялась следующим образом. Варьированием параметров θ_s , θ_c и N_c определялся такой результат обработки, в котором при совпадении ожидаемого числа кластеров $5 \leq N_c \leq 7$ с расчетным значением суммарная дисперсия распределения спектров в кластерах относительно их центров была бы минимальной. Диапазон изменения параметров определялся из предварительного анализа формы спектров и данных о влиянии состояния реактора на шумы мощности. При анализе, осуществляемом по информативным признакам спектров, процедура оптимизации была аналогична вышеописанной. Следует отметить, что в обоих случаях оптимальные результаты обработки полностью

соответствовали друг другу. Распределения спектральных оценок по кластерам в оптимальных вариантах обработки представлены в табл.2 /номера спектров, входящих в кластеры, соответствуют номерам табл.1/.

Как показали результаты обработки, частотные области наиболее информативных признаков спектров, полученных в соответствии с условиями /2/, практически не зависят от величины ϵ и Δf /в пределах изменения, отмеченных в табл.2/ и ограничены участками спектров на частотах: $f_1 = [0,0,2]$ Гц, $f_2 = 0,8$ Гц, $f_3 = 1,63$ Гц. Из дополнительных исследований известно, что низкочастотная область $f \leq 0,6$ Гц определяется условиями протекания натрия через активную зону, а колебания мощности на частотах 0,8 и 1,63 Гц однозначно связаны с поперечными вибрациями подвижных отражателей. Таким образом, формализованный отбор информативных признаков спектров по энтропии позволяет получить физически интерпретируемые результаты кластерного анализа. Кроме того, объем обрабатываемой информации уменьшается более, чем на порядок, и соответственно сокращается время обработки. Из анализа данных таблиц 1 и 2, а также сопоставления рис.1, 2 видно, что результаты кластеризации спектров практически не зависят от величины ϵ и Δf и соответствуют физическим параметрам реактора: средней мощности W и расходу натрия через активную зону G . Во всех случаях выделяются кластеры 1 /спектры 6,7/, 2 /спектры 8,9,10/, 3 /спектр 18/, 4/1,2,3,4/ и 5 /11,13,14,15,16,17/. Спектр 18 отличается от родственных / по условиям работы реактора/ спектров аномальным изменением формы в низкочастотной области, т.е. отражает изменение условий протекания натрия через активную зону, а спектр 19 отличается от соседних спектров величиной пика на частоте 1,63 Гц, что связано с некоторыми изменениями в колебаниях ПО. Надо заметить, что во всех указанных выше случаях параметры реактора находились в области допустимых отклонений и не классифицировались как аварийные. Данные табл.2 показывают также, что использование того или иного метода обработки временных рядов и связанных с этим допустимых значений ϵ и Δf зависит от цели классификации: определяются ли типы спектров, форма которых зависит только от статических параметров реактора / W, G и т.д./, или в процессе классификации необходимо идентифицировать и типичные отклонения параметров / $\Delta W, \Delta G$ и т.д./.

Если ограничиться идентификацией только W и G при условии независимости среднего уровня осевых колебаний ПО от W , что справедливо в первом приближении, то, как видно из табл.2, наиболее оптимальным является алгоритм, использующий наименьший объем данных /при $\epsilon = 18\%$ и $\Delta f = 0,024$ Гц/, который, кроме того, обладает и наибольшим быстродействием /классический вариант метода Кули-Тьюки на основе быстрого преобразования Фурье/. Этот вариант обработки обеспечивает также и поиск спектров с аномальными отклонениями формы, как, например, спектра 13.

В целом, за исключением спектров 18 и 5, которые можно интерпретировать как аномальные по форме, а спектр 18 также и по уровню наблюдаемой дисперсии /см.табл.1/, данный алгоритм обеспечивает идентификацию статических параметров реактора W, G и среднего уровня осевых колебаний ПО. Кроме того, его использование позволяет существенно увеличить число одновременно обрабатываемых спектров /до $\sim 5 \cdot 10^3$ /, что необходимо для получения статистически достоверных результатов в задаче классификации реакторных шумов.

ВЫВОДЫ

Анализ полного алгоритма обработки временных рядов в задаче классификации спектров колебаний мощности реактора ИБР-2 показал:

1. Наблюдаемые флуктуации мощности реактора допускают использование наиболее быстродействующего алгоритма получения спектральных оценок по методу Кули-Тьюки с допустимой статистической погрешностью $\epsilon = 20\%$ и разрешающей способностью $\Delta f = 0,025$ Гц.
2. Формализованный отбор информативных параметров спектров по энтропии их распределения позволяет более чем на порядок сжать объем первичной информации.
3. Звристический подход в алгоритме классификации позволяет выделить типичные структуры спектров и идентифицировать их вид с уровнем средней мощности реактора, расходом натрия через активную зону и уровнем поперечных колебаний ПО.

ЛИТЕРАТУРА

1. Gonzalez R.C., Fry D.N. and Kryter R.C. IEEE Trans. on Nuclear Science NS-21, 1, 1974, p.1.
2. Gonzalez R.C., Mowington L.C. IEEE Trans. on Systems, Man and Cybernetics, SMC-7, 10, 1977, p.717.
3. Пепельшев Ю.Н., Рогов А.Д. ОИЯИ, P13-84-73, Дубна, 1984.

Рукопись поступила в издательский отдел
5 ноября 1985 года.

Пепельшев Ю.Н. 13-85-795
Исследование алгоритма распознавания образов
в задаче классификации спектров шумов реактора ИБР-2

Рассматривается возможность использования эвристического алгоритма распознавания образов ISODATA для классификации спектров шумов мощности реактора ИБР-2. На примере обучающей выборки шумов мощности, измеренных при различных состояниях реактора, исследуются отдельные процедуры полного алгоритма распознавания, включающие способ получения спектральных оценок, формализованный алгоритм выбора информативных признаков спектров, а также возможность получения оптимальных параметров алгоритма. Показано, что оптимизация параметров, проведенная на различных этапах обработки, позволяет более чем на порядок сжать объем исходной информации и увеличить число одновременно обрабатываемых спектров до $\sim 3 \cdot 10^3$. Отмечено, что в диапазоне средней мощности реактора W от 0,16 до 1 МВт и расхода теплоносителя через активную зону G от 26 до 80 м³/ч вид спектров шумов мощности реактора ИБР-2 определяется значением W, G и интенсивностью колебаний подвижного отражателя. Указанное свойство шумов может быть использовано для диагностики состояния реактора.

Работа выполнена в Лаборатории нейтронной физики ОИЯИ.

Сообщение Объединенного института ядерных исследований. Дубна 1985

Перевод О.С.Виноградовой.

Pepelyshev Yu.N. 13-85-795
Investigation of Pattern Recognition Algorithm
to Classify the IBR-2 Reactor Noise Spectra

The possibility of application of ISODATA heuristic algorithm of pattern recognition to classify the IBR-2 noise power spectra is considered. The individual procedures of full pattern recognition algorithm, including the method of obtaining spectral evaluations, formalised algorithm of information spectra sign choice, and the possibility of getting the algorithm optimum parameters are examined. Selection of reactor power noises measured at different reactor states was taken as an example. It is shown that parameter optimization carried out at different stages of processing, permits to decrease by more than an order the volume of initial information and to increase the number of simultaneously processed spectra up to $3 \cdot 10^3$. In the range of the reactor mean power W (0.16 upto 1 MW) and coolant flow rate through the reactor core G (26 upto 80 m³ per hour) the type of IBR-2 noise power is determined by W, G values and by the intensity of the moving reflector oscillations. The property of the reactor noises indicated may be applied to the reactor state diagnosis.

The investigation has been performed at the Laboratory of Neutron Physics, JINR.

Communication of the Joint Institute for Nuclear Research, Dubna 1985