

**СООБЩЕНИЯ
ОБЪЕДИНЕННОГО
ИНСТИТУТА
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ
ДУБНА**

Н 623

P10-88-854

Н.М.Никитюк

**ТЕОРЕТИЧЕСКИЕ ОСНОВЫ СОЗДАНИЯ
СПЕЦИАЛИЗИРОВАННЫХ ПРОЦЕССОРОВ
ДЛЯ ОБРАБОТКИ СОБЫТИЙ
С КЛАСТЕРАМИ**

1988

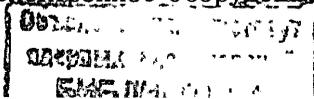
1. ПОСТАНОВКА ЗАДАЧИ

Метод синдромного кодирования, основы которого заложены в работах^{/1-4/}, эффективно используется для создания мажоритарных схем совпадений на большое число входов^{/5-8/} с широкими функциональными возможностями, параллельных шифраторов на t сигналов, где $t > 1$ ^{/9-11/}, для построения трековых и координатных процессоров^{/9-11/}. Перспективность этого метода подтверждается также работой^{/15/}, где описывается специализированный процессор, с помощью которого выполняется отбор событий по углу рассеяния частицы за 80 нс. При создании этого процессора используются алгоритмы и методы сжатия данных, предложенные нами в работах^{/1,16/}, где показано, что итеративные коды могут быть успешно использованы для сжатия данных, регистрируемых в спектрометрах физики высоких энергий. Причем имеется возможность идентифицировать топологически сложные события с кластерами. В какой-то степени в простейшем виде итеративный код использовался в процессоре^{/17/}. В алгебраической теории кодирования известны коды, с помощью которых можно эффективно исправлять как независимые ошибки, так и пакеты ошибок. Такие коды получили название кодов Рида — Соломона, и они широко используются в цифровой космической связи^{/18, 19/}, а также для коррекции ошибок в матрицах памяти с произвольной выборкой^{/20, 21/}. Цель данной работы состоит в том, чтобы показать возможность использования РС-кодов для сжатия и обработки данных, регистрируемых в многоканальных детекторах заряженных частиц, в которых кластерные события составляют значительную часть регистрируемой информации.

2. КОДЫ РИДА — СОЛОМОНА

Коды Рида — Соломона (РС-коды) являются разновидностью так называемых БЧХ-кодов, применение теории которых было рассмотрено в работах^{/1,2,4,5,8,9,14/}. Принципиальное различие заключается в том, что если с помощью БЧХ-кодов можно построить устройства для обнаружения и вычисления координат t независимых сигналов, поступающих от позиционно-чувствительного детектора, то РС-коды можно использовать для регистрации целых символов, которым можно сопоставить такие физические события, от которых генерируются кластеры.

Естественно, что для получения такого качества необходимо затратить дополнительное электронное оборудование (микросхемы для вы-



числения синдрома). Во-первых, для вычисления координат t символов требуется $2t$ проверок на четность, т.е. в два раза больше, чем это необходимо для БЧХ-кодов. Во-вторых, для вычисления синдрома РС-кода наряду со схемами проверки на четность необходимы схемы умножения двух элементов в поле Галуа $GF(2^m)$.

РС-коды имеют длину, состоящую из $n = 2^m - 1$ символов, причем каждый символ содержит m бит. Среди этих n символов имеется $2^m - 1 - 2t$ информационных и $2t$ проверочных символов, где t — количество ошибочных символов, исправляемых данным кодом. В соответствии с методом синдромного кодирования $2^m - 1 - 2t$ информационных символов рассматриваются как нулевые, а группы сигналов, поступающих от детектора, — как ошибочные символы, равные проверочным символам.

В общем виде матрица проверочных соотношений РС-кода, которая является одновременно и кодирующей матрицей^{/2,3/}, имеет вид

$$H^T = \begin{pmatrix} a^0 & a^0 & \dots & a^0 \\ a^1 & (a^1)^2 & \dots & (a^1)^t \\ a^2 & (a^2)^2 & \dots & (a^2)^t \\ \vdots & \vdots & \ddots & \vdots \\ a^{n-1} & (a^{n-1})^2 & \dots & (a^{n-1})^t \end{pmatrix} \quad (1)$$

При $t = 2$ и $m = 4$ матрица (1) имеет вид

Группы входов

*	0	a^0	a^0	a^0	a^0
*	1	a^1	a^2	a^3	a^4
*	2	a^2	a^4	a^6	a^8
	3	a^3	a^6	a^9	a^{12}
	4	a^4	a^8	a^{12}	a^1
	5	a^5	a^{10}	a^0	a^5
	6	a^6	a^{12}	a^3	a^9
	7	a^7	a^{14}	a^6	a^{13}
	8	a^8	a^1	a^9	a^2
	9	a^9	a^3	a^{12}	a^6
	10	a^{10}	a^5	a^0	a^{10}
	11	a^{11}	a^7	a^3	a^{14}
	12	a^{12}	a^9	a^6	a^9
	13	a^{13}	a^{11}	a^9	a^7
	14	a^{14}	a^{13}	a^{12}	a^{11}

В матрице (2) значения $(a^0 - a^{14})$ представляют собой элементы поля Галуа $GF(2^4)$, которые рассматриваются как символы РС-кода. Правила построения элементов поля Галуа подробно рассмотрены в работах^{/2,9,22/}. В поле Галуа $GF(2^4)$ имеется 15 ненулевых элементов: $a^0 = 1000$, $a^1 = 0100$; $a^2 = 0010$; $a^3 = 0001$; $a^4 = 1100$; $a^5 = 0110$; $a^6 = 0011$; $a^7 = 1101$; $a^8 = 1010$; $a^9 = 0101$; $a^{10} = 1110$; $a^{11} = 0111$; $a^{12} = 1111$; $a^{13} = 1011$ и $a^{14} = 1001$.

Поскольку каналы регистрации нумеруются степенями элементов поля Галуа $GF(2^4)$, то вектор координат сработавших датчиков $e(X)$ можно задать перечнем значений его ненулевых компонент и позиций, на которых они расположены. Эти позиции определяются номерами позиций ошибок. Например, если символ появился в третьей группе, то его позиция равна a^2 .

Далее, каждая ненулевая компонента вектора координат сработавших датчиков определяется парой элементов Y_i (значением символа или кластера) и X_i — координатой кластера. Если было t кластерных событий, то вектор $e(X)$ имеет t ненулевых компонент, и поэтому для описания таких событий требуется t пар (X_i, Y_i) . Величины S_i формируются в соответствии с равенствами

$$S_j = \sum Y_i X_i^j, \quad (3)$$

где Y_i — символы, поступающие на вход схемы для вычисления синдрома (коды событий).

Следует отметить, что в более простом случае, когда используется теория БЧХ-кодов^{/2,5,9,14/}, значения $Y_i = 0$ или 1, поэтому уравнения (3) имеют вид

$$S_j = \sum X_i^j, \quad 1 \leq j \leq 2t,$$

причем для вычисления кода синдрома используются только нечетные симметрические функции $S_1, S_3, S_5 \dots S_{2t-1}$.

Уравнения (3) являются нелинейными, и поэтому при $t > 5$ их непосредственное решение затрудняется. При $t \leq 5$ известны табличные методы решения таких уравнений, которые аналогичны методам, используемым в теории БЧХ-кодов^{/20,21/}. Эти методы, естественно, представляют наибольший интерес, поскольку определение значений X_i и Y_i может быть выполнено относительно быстро.

В технике кодирования и в космической связи используются более экономичные последовательные кодирующие устройства, основанные на алгоритме декодирования, предложенном У.Питерсоном^{/18,19,23/}. В технике ядерного эксперимента последовательные методы кодирования и декодирования могут найти применение при обработке данных в режиме оф-лайн. Поскольку последовательный метод Питерсона является универсальным, т.е. позволяет находить значения X_i и Y_i в общем виде, то рассмотрим кратко его суть.

Известно, что значения X_i удовлетворяют так называемому ключевому уравнению

$$X^t + \sigma_1 X^{t-1} + \sigma_2 X^{t-2} + \dots + \sigma_{t-1} X + \sigma_t = 0, \quad (5)$$

где $\sigma_1, \sigma_2, \dots, \sigma_t$ — элементарные симметрические функции от X_i , которые связаны с S_j формулами Ньютона

$$S_j \sigma_t + S_{j+1} \sigma_{t-1} + \dots + S_{j+t} \sigma_1 + S_{j+t} = 0. \quad (6)$$

В результате алгоритм нахождения X_i и Y_i состоит из нескольких этапов.

1. Вычисление синдрома S_1, S_2, \dots, S_{2t} , которое сводится к проверке на четность.

2. Вычисление $\sigma_1, \sigma_2, \dots, \sigma_t$ по S_j .

3. Вычисление координат X_i по σ_i .

4. Вычисление значений кластеров Y_i по X_i и S_j из (3).

Для определения X_i по σ_i необходимо в уравнении (5) последовательно подставить значение всех элементов поля Галуа $GF(2^m)$ и те из них, которые, обращают в нуль уравнение (5) и являются искомыми значениями X_i . Рассматривая применение теории РС-кодов к задаче регистрации кластерных событий в спектрометрах физики высоких энергий, прежде всего необходимо найти метод определения величины множественности t . Алгоритм нахождения величины t следует из теоремы: матрица

$$L_t = \begin{vmatrix} S_1 & S_2 & S_3 & \dots & S_t \\ S_2 & S_3 & S_4 & \dots & S_{t+1} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ S_t & S_{t+1} & S_{t+2} & \dots & S_{2t+1} \end{vmatrix} \quad (7)$$

невыврождена, если величины S_j образуются точно из t различных ненулевых пар (X, Y_i) , и матрица (7) вырождена, если S_j образуются из меньшего чем t числа ненулевых пар (X_i, Y_i) . Доказательство этой теоремы дано в [23, с.317].

Таким образом, анализируя значения детерминанта матрицы L_t , можно точно определить величину множественности кластерных событий. Другими словами, свойства $\det L_t$ таковы, что если, например, $t = 1$, то $\det L_1 \neq 0$, но все остальные определители $\det L_2, \det L_3$ и т.д. равны нулю. Если же $t = 2$, то $\det L_1 \neq 0, \det L_2 \neq 0$, но $\det L_3, \det L_4$ и т.д. равны нулю.

При $t = 3$ имеем

$$L_t = \begin{vmatrix} S_1 & S_2 & S_3 \\ S_2 & S_3 & S_4 \\ S_3 & S_4 & S_5 \end{vmatrix}. \quad (8)$$

Нетрудно подсчитать, что

$$\det L_1 = S_1, \quad \det L_2 = S_1 S_3 + S_2^2 \quad \text{и} \quad (9)$$

$$\det L_3 = S_1 S_3 S_5 + S_1 S_4^2 + S_2^2 S_5 + S_3^3.$$

Следует отметить, что в поле Галуа $GF(2^m)$ операции сложения и вычитания равносильны и выполняются по модулю два. Поэтому приведение подобных членов выполняется по правилу: четное число одинаковых членов равно нулю, а нечетное число — одному члену. Например, $S_1 + S_1 = 0$, но $S_1 + S_1 + S_1 = S_1$.

Рассмотрим конкретный пример. Допустим, что имеется многоканальный детектор, содержащий $n = 2^4 - 1 = 15$ групп выходов, причем каждая группа содержит по $m = 4$ каналов регистрации. Предположим, что длина кластера не превышает величины m , причем под кластерами будем понимать такие значения, как 1111, 1001, 1101 и т.д. (равные элементам поля Галуа $GF(2^4)$). В зависимости от постановки задачи такие символы, как 1001, можно рассматривать и как независимые срабатывания двух датчиков детектора от двух различных частей.

Допустим, что в детекторе зарегистрировано $t = 2$ кластеров на позициях $X_1 = a^0$ и $X_2 = a^2$, причем $Y_1 = a^7 = 1101$ и $Y_2 = a^{11} = 0111$. Эти позиции отмечены знаком * в (2). Метод вычисления синдрома поясняется также с помощью схемы, приведенной на рисунке. Здесь приняты следующие обозначения: 0-15 — группа выходов многоканального детектора, 16-30 — группа усилителей-формирователей, 31-90 — группа схем умножения двух элементов в поле Галуа $GF(2^4)$, 91 — группа схем проверки на четность, 92, 93 — ППЗУ. Группы входов пронумерованы также элементами поля Галуа $a^0 \div a^{14}$. Рассмотрим отдельно три случая $t = 1, t = 2$ и $t = 3$.

1. Допустим, что кластер $a^{11} = 0111$ поступил на третью группу входов, тогда из (4) имеем

$$\begin{aligned} S_1 &= a^{11} a^2 = a^{13}, \\ S_2 &= a^{11} a^4 = a^{15} = a^0, \\ S_3 &= a^{11} a^6 = a^{17} = a^2, \\ S_4 &= a^{11} a^8 = a^{19} = a^4. \end{aligned} \quad (10)$$

Из соотношений (9) вычисляем значения определителей

$$\begin{aligned} \det L_1 &= S_1 = a^{13} \neq 0, \\ \det L_2 &= S_1 S_3 + S_2^2 = a^{13} a^2 + a^0 = 0, \\ \det L_3 &= a^{13} a^2 a^6 + a^{13} a^8 + a^0 a^6 + a^6 = 0. \end{aligned}$$

Допустим, что кластер $a^7 = 1101$ поступил на первую группу входов, а на третью группу входов поступил такой же кластер, как и в первом случае, тогда имеем для $t = 2$.

$$\begin{aligned} S_1 &= a^7 a^0 + a^{11} a^2 = a^5, \\ S_2 &= a^7 a^0 + a^{11} a^4 = a^9, \\ S_3 &= a^7 a^0 + a^{11} a^6 = a^{12}, \\ S_4 &= a^7 a^0 + a^{11} a^8 = a^3. \end{aligned} \quad (11)$$

$$\sigma_1 = \frac{S_3 S_2^2 + S_1 S_2 S_4}{S_2^3 + S_1 S_2 S_3}, \quad (12)$$

$$\sigma_2 = \frac{S_2 S_4 + S_3^2}{S_2^2 + S_1 S_3}; \quad (13)$$

Из равенств (12, 13) находим $\sigma_1 = a^8$ и $\sigma_2 = a^2$, тогда уравнение (12) принимает вид

$$X^2 + a^8 X + a^2 = 0. \quad (14)$$

Можно проверить, что при $X_1 = a^0$ и $X_2 = a^2$ уравнение (14) обращается в тождество:

$$(a^0)^2 + a^8 a^0 + a^2 = a^0 + a^8 + a^2 = 0,$$

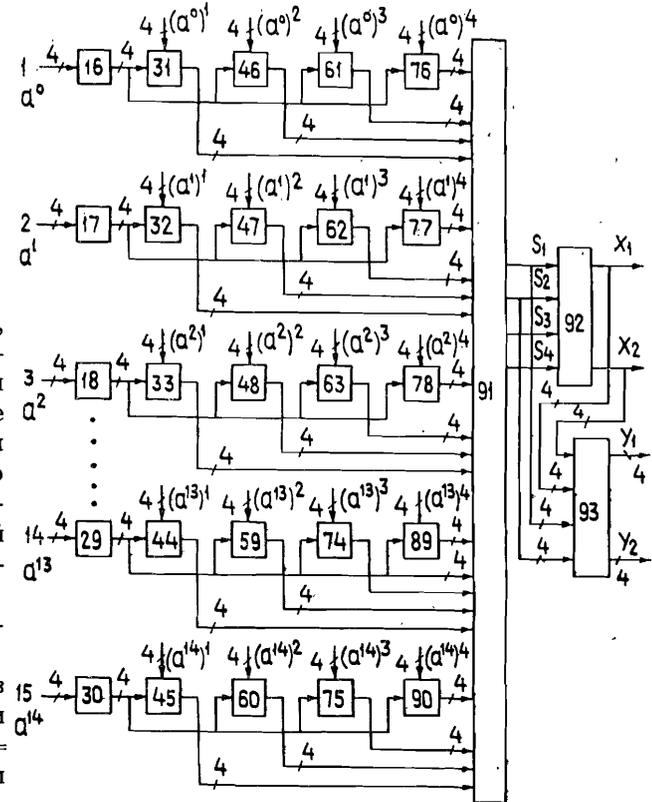
$$\begin{array}{r} 1000 \\ + 1010 \\ + 0010 \\ \hline 0000 \end{array}$$

Аналогично,

$$(a^2)^2 + a^8 a^2 + a^2 = 0,$$

$$\begin{array}{r} 1100 \\ + 1110 \\ 0010 \\ \hline 0000 \end{array}$$

Блок-схема для вычисления кода синдрома при $t = 2$.



Следует подчеркнуть, что значения X_1 получаются не в обыкновенном двоичном коде, а в виде значений элементов поля Галуа $GF(2^m)$. Из этого примера становится ясным смысл обозначений групп входов элементов поля на рисунке.

Можно проверить также, что уравнения (13) и (14) обращаются в тождество при $Y_1 = a^7$ и $Y_2 = a^{11}$. С учетом $X_1 = a^0$ и $X_2 = a^2$ имеем

$$Y_1 a^0 + Y_2 a^2 = a^5,$$

$$Y_1 (a^0)^2 + Y_2 (a^2)^2 = a^9,$$

$$\text{и далее } a^7 a^0 + a^{11} a^2 = a^5,$$

$$\text{так как } a^7 + a^{13} = a^5.$$

Аналогично,

$$a^7 a^0 + a^{11} a^4 = a^9.$$

На конкретном примере мы рассмотрели использование метода последовательного декодирования, т.е. нахождение значений X_1 и Y_1 , хотя существуют более быстрые методы декодирования РС-кодов [20, 21]. Применительно к использованию метода синдромного кодирования для регистрации кластерных событий в спектрометрах физики высоких энергий с целью получения высокого быстродействия имеет смысл использовать метод прямого декодирования кода синдрома S_1, S_2, \dots, S_{2t} . Для этого вначале по значению синдрома определяются X_1 , а затем, декодируя X_1 и S_j , можно определить Y_1 . На рисунке приведен пример прямого декодирования синдрома для $t = 2$.

В настоящее время промышленностью освоен выпуск ПЗУ емкостью 256К и выше с временем доступа менее порядка 100 нс. Использование таких ПЗУ позволит создавать быстродействующие процессоры для обработки событий с кластерами в реальном времени.

Был рассмотрен конкретный пример, когда длина кластера $t = 4$. Если же по условию задачи требуется регистрация кластеров длиной $t = 6$, необходимо построить поле Галуа $GF(2^6)$. Тогда число групп каналов регистрации $n = 2^6 - 1 = 63$. В остальном расчетные соотношения остаются такими же, как и для $m = 4$.

ЛИТЕРАТУРА

1. Nikitjuk N.M., Radzhabov R.S., Schafranov M.D. — NIM, 1978, v.155, No.1, p.485.
2. Никитюк Н.М. — ОИЯИ, P11-80-484, Дубна, 1980, 15с.
3. Никитюк Н.М. — ПТЭ, №3, 1983, с.74.
4. Nikitjuk N.M. — JINR, E11-87-10, Dubna, 1987.
5. Nikitjuk N.M. — JINR, E10-88-28, Dubna, 1987.
6. Никитюк Н.М. — ПТЭ, 1987, №3, с.59.
7. Никитюк Н.М., Селиков А.Н. — ПТЭ, 1987, №6, с.
8. Гайдамака Р.И. и др. — ОИЯИ, 13-82-628, Дубна, 1982.
9. Никитюк Н.М., Раджабов Р.С., Шафранов М.Д. — ПТЭ, 1978, №4, с.95.
10. Никитюк Н.М. — ПТЭ, 1986, №6, с.77.
11. Комолов Л.Н. и др. — ПТЭ, 1987, №6, с.
12. Калининков В.А., Крастев В.Р., Чудаков Е.А. — ПТЭ, 1986, №3, с.105.
13. Гуськов Б.Н., Иванченко И.М., Калининков В.А. — ОИЯИ, 13-87-49, Дубна, 1987.
14. Никитюк Н.М. — ОИЯИ, P10-87-254, Дубна, 1987.
15. Gustafsson L., Hagberg E. — NIM, 1988, No.2, p.521.
16. Никитюк Н.М. — ОИЯИ, P10-87-266, Дубна, 1987.
17. Kaschuk A.P., Golovtsov V.L. — LNPI, No.395, Leningrad, 1978.
18. Liu K.J. — IEEE Trans. on Comput., 1984, v.C-33, No.2, p.178.
19. Heu In-Shek, Reed I.S., Truong T.K. — IEEE Trans. on Comput., 1984, v.C-33, No.10, p.906.
20. Okano H., Imai H. — IEEE Trans. on Comput., 1987, v.C-36, No.10, p.1165.
21. Deng R.H., Costello D.I. — IEEE. Trans. on Comput., 1987, v.C-36, No.11, p.1359.
22. Никитюк Н.М. — От современной алгебры к специализированным процессорам. Лекции для молодых ученых. ОИЯИ, P10-87-401, Дубна, 1987.
23. Питерсон У., Уэлдон Э. Исправляющие ошибки. М.: Мир, 1976, с.315.

Рукопись поступила в издательский отдел
12 декабря 1988 года.

Никитюк Н.М.

P10-88-854

Теоретические основы создания специализированных процессоров для обработки событий с кластерами

Рассмотрены теоретические основы создания специализированных процессоров для быстрой регистрации и обработки кластерных событий. Работа процессоров основана на хорошо разработанных методах кодирования и декодирования так называемых кодов Рида — Соломона, которые широко используются в космической цифровой связи и в матрицах памяти с байтовой структурой для коррекции ошибок. Приводятся расчеты, необходимые для построения устройств как для определения количества зарегистрированных кластеров, так и для получения данных о координатах и образах (значениях) кластерных событий.

Работа выполнена в Лаборатории высоких энергий ОИЯИ.

Сообщение Объединенного института ядерных исследований. Дубна 1988

Перевод О.С.Виноградовой

Nikitjuk N.M.

P10-88-854

Theoretical Foundations of Creation of Special Purpose Processors for Processing Events with Clusters

Theoretical foundations of creating the special purpose processors for fast registration and processing of cluster events are considered. Processor operation is based on well developed methods for coding and decoding the so-called Reed-Solomon codes widely used in cosmic digital link and in memory matrix with byte structure for mistake correction. Calculations needed for constructing the devices both for determining the quantity of registered clusters and for deriving the data on coordinates and images (values) of cluster events are presented.

The investigation has been performed at the Laboratory of High Energies, JINR.

Communication of the Joint Institute for Nuclear Research. Dubna 1988