

РОССИЙСКАЯ АКАДЕМИЯ НАУК

ФИЗИЧЕСКИЙ
ИНСТИТУТ
имени
П.Н.Лебедева



Ф И А Н

ПРЕПРИНТ

Н.И. СТАРКОВ

25

**АНАЛИЗ ИЗОБРАЖЕНИЙ
RICH-ДЕТЕКТОРА С ПОМОЩЬЮ
НЕЙРОННОЙ СЕТИ**

МОСКВА 2003

АННОТАЦИЯ

В работе рассмотрена возможность обработки сигналов от RICH-детектора с помощью нейронной сети при прохождении через него частицы высокой энергии под углом к чувствительной плоскости. Показано, что в этом случае для определения заряда и скорости частицы требуется намного меньшая информация, чем при использовании аналитических методов обработки сигнала. Кроме того выявлена сильная корреляция между углом частицы θ и проекцией на чувствительную область точки ее входа (x_0, y_0) в радиатор детектора. Это дает возможность по информации об одной из этих характеристик определить вторую.

ABSTRACT

In this article the possibility of the processing of RICH-detector signals from a high energy particle by a neural network is considered. It is shown that in this case to determine a charge and velocity of particle it is necessary less information in comparison with an analytic method. Besides the large correlation between the particle angle θ and the projection of an entry point into radiator (x_0, y_0) is detected. This gives the possibility knowing one of the last characteristics to determine of other.

Анализ изображений RICH-детектора с помощью нейронной сети.

Н.И. Старков

Физический институт им. П.Н. Лебедева РАН РФ

starkov@sci.lebedev.ru

Введение.

В работах [1,2] описана конструкция и характеристики RICH-детектора, предназначенного для изучения свойств заряженных частиц космического излучения в составе магнитного спектрометра. Основной задачей данного устройства было определение заряда q и скорости β частицы, прошедшей через установку. Алгоритм, использовавшийся для выполнения этой задачи, помимо знания координат сработавших фотоэлектронных умножителей (ФЭУ) требовал наличия дополнительной информации от магнитного спектрометра: координат x_0 и y_0 точки входа частицы в детектор и направление ее движения (азимутального φ и полярного углов θ).

В последнее десятилетие широкое применение в физике высоких энергий и элементарных частиц нашли нейронные сети. С их помощью разделяются кварк-глюонные струи [3], идентифицируются частицы в конечном состоянии [4], осуществляется режекция нейтронов и гамма-квантов [5], реконструируется продольное фазовое пространство [6] и т. д. В настоящей работе будет показано, что использование нейронной сети позволяет существенно сократить число необходимых данных при анализе изображений RICH-детектора. В частности, для определения заряда и скорости достаточно знать положение сработавших ФЭУ, полную амплитуду сигнала от них и азимутальный угол φ .

Модельный вариант детектора содержит радиатор NaF толщиной 1 см. и матрицу фотодетекторов площадью 1 м^2 , расположенные параллельно друг другу на расстоянии 10 см [2]. Матрица собрана из более чем 1100 ФЭУ с диаметром фотокатода 25 мм и внешним диаметром 32 мм. Входные окна ФЭУ изготовлены из стекла, что приводит к ограничению фотонов, регистрируемых RICH-детектором, диапазоном длин волн 300-650 нм. В этом диапазоне показатель преломления NaF составляет $n \approx 1,334$, а его дисперсия менее 1%.

Релятивистская заряженная частица, проходя через радиатор со скоростью $v = \beta c > c \cdot n^{-1}$, излучает фотоны под углом $\gamma = \arccos(1/\beta n)$ к ее траектории. Интенсивность излучения прямо пропорциональна квадрату заряда частицы и приблизительно обратно пропорциональна квадрату длины волны. На границе радиатора фотоны испытывают преломление и частичное отражение, в результате чего на матрице ФЭУ ими образуется фигура определенной формы. При углах частицы θ , близких к перпендикуляру к плоскости матрицы ФЭУ ($\theta < 8^\circ \div 12^\circ$, в зависимости от скорости), эта фигура замкнутая (квазиэллипс), при больших – разомкнутая, вследствие полного отражения на границе радиатора. Часть фотонов, попавших в область чувствительности ФЭУ, вызывает выбивание фотоэлектронов и его

срабатывание. В работе [2] приведено описание программы для розыгрыша методом Монте-Карло событий прохождения заряженной частицы через RICH-детектор. В ней учитывались основные неопределенности, влияющие на точность конечного результата: поглощение, отражение и преломление света в радиаторе и окнах ФЭУ, спектр черенковских фотонов, спектральная чувствительность ФЭУ. На рисунках 1а и 1б приведены несколько типичных фигур, образуемых иницированными ФЭУ, при различных значениях скорости частицы β и угла падения θ . При больших β они занимают довольно значительную площадь из-за дисперсии радиатора и эффектов преломления. Кроме того, при больших углах падения частицы ($\theta > 8^\circ \div 12^\circ$) фигура оказывается незамкнутой в результате полного внутреннего отражения части излучения.

Результаты моделирования [2] продемонстрировали возможность восстановления скорости (с точностью $\sim 0,3\%$) и заряда частицы. Для этого необходимы координаты иницированных ФЭУ, полная амплитуда снимаемого сигнала, а также геометрические характеристики траектории частицы: координаты ее точки входа в радиатор и углы θ и φ . Такие жесткие требования обеспечить довольно трудно, поэтому были предприняты поиски методов, позволяющих ослабить их. Одним из возможных вариантов является использование нейронной сети.

Принцип работы НС.

В данной работе была использована многослойная нейронная сеть – многоуровневый персептрон, с обратным распространением ошибки. Она используется для установления нелинейного многопараметрического отображения одного многомерного пространства векторов (вектора входа V_{in} , пространство признаков) на другое (вектора выхода V_{out} , пространство значений) с помощью обучающего набора. Это означает, что, будучи подготовленная к работе, при подаче на вход сети одного из векторов V_{in} , она должна дать на выходе соответствующий вектор V_{out} .

Устройство и принцип работы нейронных сетей описаны во многих работах (см., например, [7]). Поэтому здесь мы приведем только упрощенную схему действия персептрона с обратным распространением ошибки.

Нейронная сеть такого типа состоит из трех или большего числа последовательных слоев, каждый из которых включает некоторое количество нейронов. Первый и последний слои являются входом и выходом, соответственно, поэтому количество нейронов в них должны быть равны размерностям соответствующих пространств. Промежуточные слои носят название скрытые. Каждый нейрон одного слоя имеет связь с каждым нейроном предыдущего и последующего слоев. Отдельный нейрон характеризуется числовым значением u (которое, как правило, входит в интервал $[0, 1]$), а каждая связь своим весом w . Нейроны первого слоя имеют значения, равные компонентам входного вектора. Значения нейронов

последующих слоев являются нелинейной функцией значений нейронов предыдущего слоя и соответствующих весов.

Работа сети протекает в два этапа. На первом из них происходит ее обучение. Для этого с помощью экспериментов или их программных имитаторов формируется обучающее множество $\{V_{in}; V_{out}\}$ которое представляет собой набор пар входного и соответствующего выходного векторов. Размер обучающего множества (количество пар) зависит от сложности анализируемого отображения и общего числа слоев и нейронов на них. Входные вектора включают непосредственно наблюдаемые величины (показания приборов или их комбинации), а выходные – физические характеристики процесса, которые требуется определить в результате анализа (типы частиц, их заряды, скорости и т.д.). В начале обучения значения w НС иницируются небольшими случайными значениями и на ее вход по очереди подаются значения входных векторов из обучающего множества. Порядок следования векторов определяется случайным образом, чтобы предотвратить "привыкания" сети к их фиксированной последовательности. С помощью нелинейного преобразования значений u_j каждого из нейронов входного слоя и величин w_{ji} формируются значения u_i нейронов второго слоя. Затем эти значения используются для вычисления третьего слоя. Такая процедура повторяется для всех слоев вплоть до последнего - выходного. Значения нейронов последнего слоя сравниваются с соответствующими значениями выходного вектора из обучающего множества и вычисляется ошибка, например, сумма квадратов разности или сумма модулей разности. По величине ошибки с помощью специального алгоритма вычисляются поправки к значениям весов связей последнего и предпоследнего слоев. По этим поправкам последовательно вычисляются поправки к весам более ранних слоев вплоть до первого. Затем происходит подача на вход следующего входного вектора и процедура уточнения весов повторяется до получения приемлемого результата. Процедура обучения, охватывающая весь обучающий набор, принято называть эпохой. По мере обучения ошибка, как правило, уменьшается и процесс сходится, но весь цикл обучения может занять не один десяток, а то и сотни эпох. Успех обучения зависит от целого ряда причин, в том числе от удачного выбора пространства признаков и достаточном объеме обучающего множества. Упрощая, можно сказать, что процесс обучения является подгонкой процедурой для описания многопараметрической функцией (параметрами являются величины w_{ij}) упомянутого выше отображения пространства входных векторов на пространство выходных.

На эффективность работы и сходимости процесса обучения НС большое влияние оказывает правильный выбор входных сигналов. Они должны быть такими, чтобы в многомерном пространстве – пространстве признаков, образуемом ими, области с различными значениями выходных векторов разделялись как можно лучше.

Второй этап представляет собой непосредственную работу обученной НС по анализу физической информации.

Применение НС при анализе данных RICH-детектора

Исходная информация RICH-детектора, из которой можно построить пространство входных векторов, включает количество инициированных ФЭУ, их координаты и амплитуду сигнала, снимаемого с них. Теоретически возможен вариант установки, когда доступна информация об амплитуде сигнала от каждого ФЭУ, но он сложен в технической реализации и здесь рассматриваться не будет.

При обсуждении возможностей нейронной сети ограничимся рассмотрением углов, меньших 8° , при которых фигура на матрице ФЭУ остается замкнутой. На самом деле НС можно применить и в случае незамкнутой фигуры при небольшой модификации, связанной с описанием ее геометрии.

Фигура, создаваемая черенковским светом на светочувствительной матрице симметрична относительно проекции траектории на плоскость матрицы. Это свойство можно было бы использовать для идентификации азимутального угла. Однако, не все ФЭУ, на которые падают фотоны, инициируются. Кроме того, рассматриваемая конструкция обладает большой дискретностью чувствительной области, которая занимает только 55% площади детектора, что приводит к ломаному контуру фигуры и непостоянству ее формы. Особенно сильно это эффект проявляется при малых скоростях v . Вследствие этого угол φ определяется с большой погрешностью, которая сильно снижает надежность определения остальных характеристик частицы. По этой причине будем считать, что угол φ известен.

Анализ свойств RICH-детектора показал, что для определения величины заряда (вплоть до величин $q=38-40$) недостаточно знания полной амплитуды снимаемого сигнала N_a или полного количества инициированных ФЭУ N_{pmt} , так как эти величины могут быть одинаковыми у частиц с разными зарядами при разных скоростях. Кроме того, при большом количестве черенковских фотонов, то есть при больших зарядах и скоростях, вероятность срабатывания ФЭУ близка к единице, что приводит к насыщению величины N_{pmt} и снижению чувствительности RICH-детектора к заряду.

К величине скорости частицы и наклону траектории θ очень чувствительна геометрическая форма фигуры, создаваемой черенковскими фотонами. Наиболее полная информация о ней содержится в координатах положения инициированных ФЭУ. Однако для определения скорости эта информация избыточна. Анализ показал, что величины геометрического размера фигуры содержат достаточно полную информацию. В итоге в качестве основного параметра, определяющего скорость, был выбран максимальный размер фигуры Y_{max} в направлении перпендикулярном проекции траектории частицы ("малая ось" квазиэллипса) на плоскость радиатора. Он может быть найден из координат инициированных ФЭУ и угла φ . Другим возможным параметром может служить величина максимального размера фигуры X_{max} в направлении параллельном проекции траектории частицы ("большая ось" квазиэллипса). Этот параметр

чувствителен, кроме того, к величине угла θ , однако, он испытывает сильный разброс вследствие дискретности чувствительной области и поэтому фиксирует угол с большой ошибкой.

В итоге, после проведенного анализа в качестве входного вектора был выбран набор (N_a, Y_{\max}) , а в качестве выходного – набор (q, β) . На рисунке 2 показаны распределения величины N_a , а на рисунке 3 - Y_{\max} , от скорости, полученные в результате розыгрыша 2400 событий для трех ядер (${}^4\text{He}$, ${}^9\text{Be}$, ${}^{12}\text{C}$) при 10 значениях скорости и 8 значениях угла θ . Видно, что в пространстве (N_a, Y_{\max}) заряды и скорости довольно хорошо пространственно разделены. Это дало основание надеяться на возможность решения поставленной задачи с помощью НС.

Для имитации работы НС была создана программа, реализующая алгоритм обратного распространения ошибки. В качестве обучающего множества был взят упомянутый выше набор из 2400 событий. В процессе выбора оптимальной конфигурации НС рассматривались варианты с различным числом скрытых слоев и нейронов на них. Увеличение их числа приводит к росту времени обучения и требует большего объема обучающего множества, а уменьшение ухудшает качество описания. Окончательный вариант НС включал три скрытых слоя по 10 нейронов на каждом. Обучение по 2400 образцам потребовало около 2000 эпох.

Для проверки качества работы НС был сгенерирован тестовый набор входных и выходных векторов, включающий 840 событий для тех же ядер, но с промежуточными значениями скорости и угла θ . Распределения полученных величин приведены на рисунке 4 вместе с исходными значениями, обозначенными треугольником. Видно, что заряды частиц, будучи дискретными, определяются практически с абсолютной точностью. Точность определения скорости частиц колеблется в интервале от 0,1% до 0,4%. На рисунке 5 приведено распределение в одной точке ($q=2, \beta=0,933$), где разброс найденных значений β' наибольший: $\beta'=0,934\pm 0,004$. Добавление X_{\max} к набору входных параметров не улучшает описания.

В рассмотренном выше варианте совсем не участвовала информация о полярном угле θ и точке входа (x_0, y_0) частицы. Дополнительный анализ показал, что эти две характеристики очень сильно коррелируют друг с другом. В связи с этим были рассмотрены варианты работы НС с расширенными наборами входных и выходных векторов. Было использовано то обстоятельство, что цент масс инициированных ФЭУ с большой точностью попадает на проекцию трека на чувствительную плоскость и в качестве дополнительного параметра можно выбрать расстояние от центра масс ФЭУ до проекции точки входа частицы в радиатор R_0 . Зная угол φ , R_0 и координаты ФЭУ можно легко восстановить координаты точки входа частицы x_0 и y_0 . В результате, были выбраны два новых набора входных и выходных векторов: $\{(N_a, X_{\max}, Y_{\max}, \theta); (q, \beta, R_0)\}$ и $\{(N_a, X_{\max}, Y_{\max}, R_0); (q, \beta, \theta)\}$. Параметр X_{\max} был добавлен в набор входных векторов, т.к. он также чувствителен к углу θ и точке входа и делает их определение более точным.

Проверка возможностей этих наборов производилась в той же конфигурации сети и с тем же обучающим множеством, включающим 2400 событий. Проверка результативности обучения на тестирующем множестве из 840 событий показал высокую эффективность определения величин выходного вектора и в этом случае. Так точность определения угла θ составляет около 1° , а точность определения R_0 – от 3 до 5 миллиметров. Кроме того, точность определения скорости β в этом случае повышается.

В последнем из рассмотренных варианте предполагалось знание точного значения величин входного вектора, в том числе угла θ или R_0 . Может случиться так, что имеется информация об обеих этих величинах, но эта информация не точная, а с некоторой ошибкой. Учитывая сильную корреляцию θ и R_0 можно надеяться, что если во входной вектор включить эти, пусть даже и не точные, значения, с помощью нейронной сети можно будет уменьшить эту неопределенность на выходе. Более чувствительный вариант НС мог бы включать детальную информацию о положениях ФЭУ. Возможности таких вариантов и их эффективность будут рассмотрены в следующей публикации.

Литература.

1. К.В. Александров и др., "Эффективность регистрации черенковских фотонов рич-детектором с радиаторами из LiF и MgF₂, Письма в ЖТФ, 1996, т. 22, вып. 18, стр. 11-15.
2. К.В. Александров и др., "Моделирование RICH детектора AMS и реконструкция кинематических параметров регистрируемых частиц", Математическое моделирование, 1997, т. 9, № 11, стр. 33-45.
3. M.F.A. Graham, L.M. Jones, S. Herbin, "Neural network classification of quark and gluon jets", Phys. Rev., 1995, v. 51, p. 4789.
4. R. Bellotti et al., "A neural network for positron identification by transition radiation detector", Nucl. Instr. and Meth., 1994, v. A350, p. 556.
5. P. Pavlopoulos et al., "Identification of $\bar{p}p \rightarrow K^\pm \pi^\pm K^0(\bar{K}^0)$ events using artificial neural network", Nucl. Instr. and Meth., 1995, v. A359, p. 566.
6. J. Koga, T. Takeda, "Neural network reconstruction of longitudinal beam phase space from the synchrotron radiation spectrum", Nucl. Instr. and Meth., 1995, v. A363, p. 580.
7. R. Callan, "The Essence of Neural Network", Prentice Hall Europe, 1999. Laurene Fauset, "Fundamentals of Neural Networks", Prentice Hall, Upper Saddle River, New Jersey, 1994. J. Hertz, A. Krogh, R.G. Palmer, "Introduction to the Theory of Neural Computation", Addison-Wesley, Redwood City, Ca., 1991. P.C. Rognvaldsson, "An Introduction to Artificial Neural Networks", CERN School of Computing, 1991.

Подписи к рисункам.

- Рисунок 1. Изображение на чувствительной плоскости RICH-детектора, созданное засвеченными (пустые кружки) и иницированными (затемненные кружки) ФЭУ при прохождении ядра ${}^4\text{He}$ через установку.
а. $\beta=0,85$ $\theta=0^\circ$ (1), $\beta=0,97$ $\theta=12^\circ$ (2);
б. $\beta=0,97$ $\theta=8^\circ$.
Угол $\varphi=45^\circ$.
- Рисунок 2. Зависимость от скорости частицы величины сигнала от иницированных ФЭУ для трех ядер. Приведены точки обучающего множества, всего 2400 событий: 3 ядра, 10 скоростей, 8 углов. По 10 событий при каждом значении.
- Рисунок 3. Зависимость от скорости частицы величины Y_{\max} фигуры, создаваемой иницированными ФЭУ. Приведены точки обучающего множества (см. рис. 2).
- Рисунок 4. Распределение величин заряда и скорости, полученных в результате работы нейронной сети с тестовым набором (3 ядра, 7 скоростей, 4 угла, по 10 событий в каждой точке).
- Рисунок 5. Распределение скоростей, полученных в результате работы нейронной сети с тестовым набором, в точке наибольшего разброса скорости: $\beta=0,933\pm 0,004$. Истинное значение скорости $\beta=0,933$.

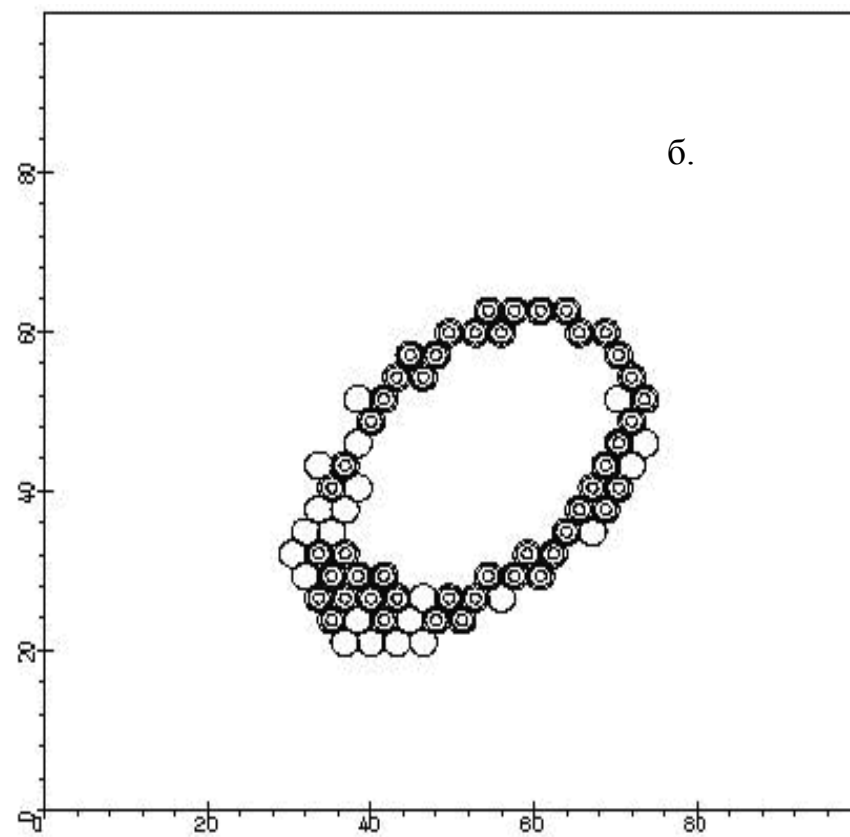
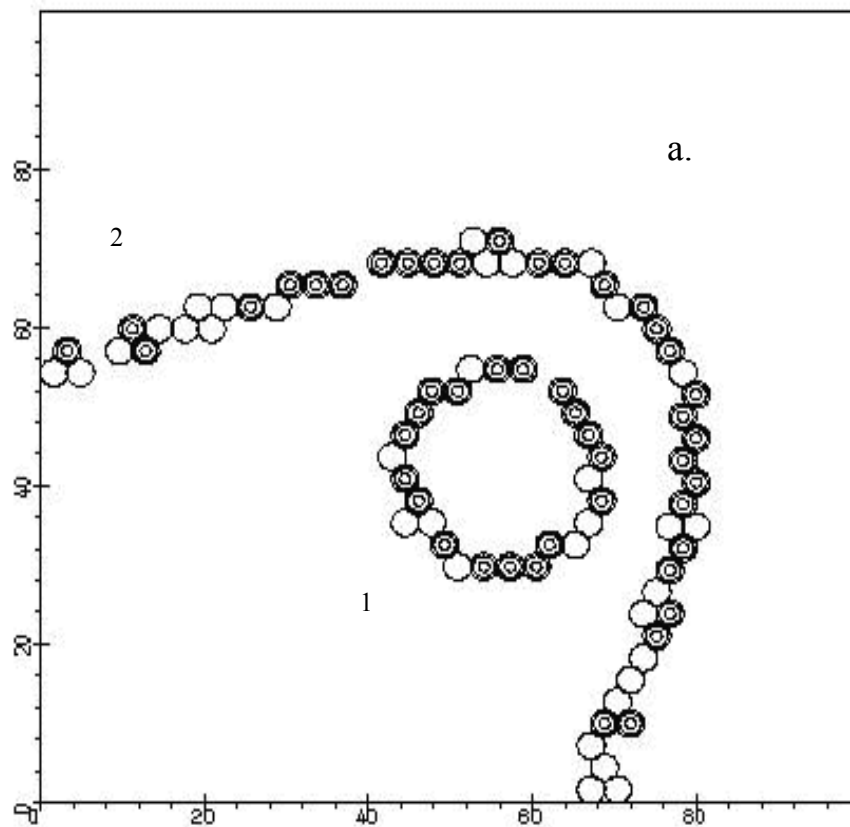


Рисунок 1.

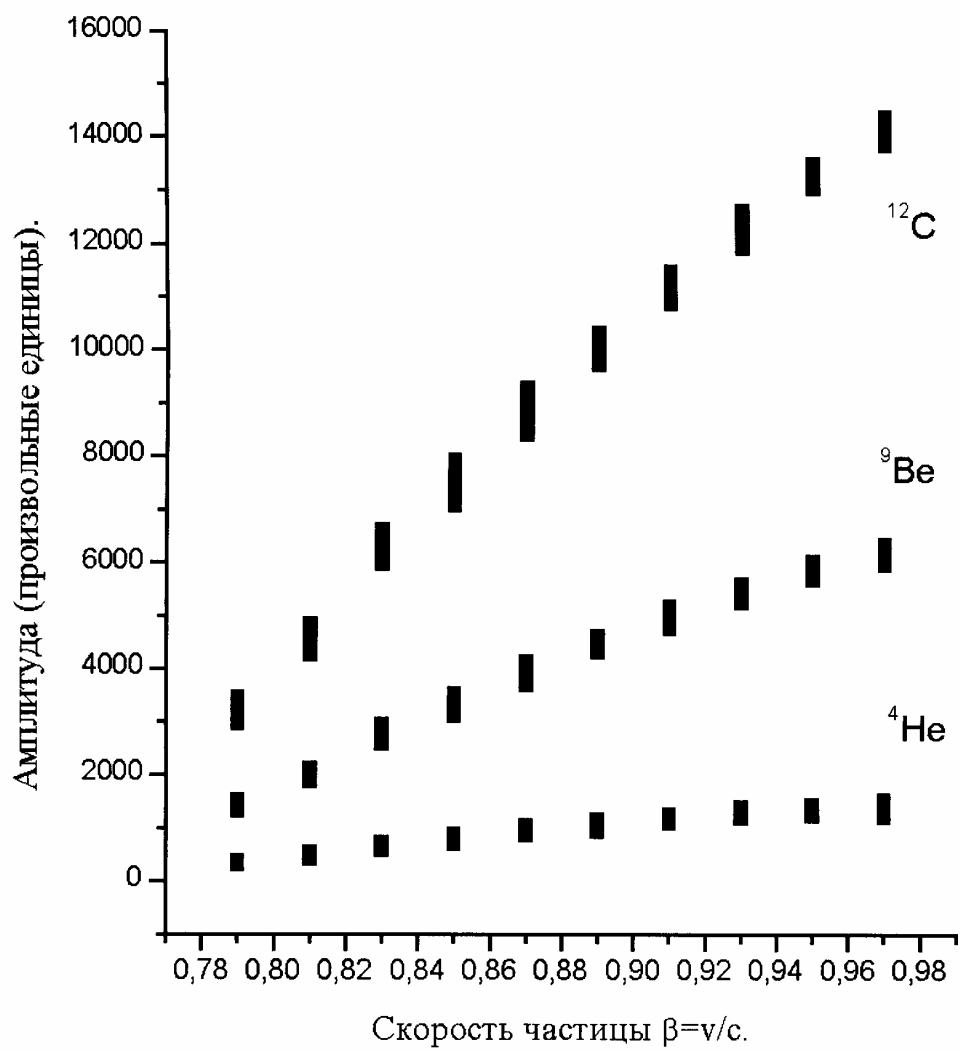


Рисунок 2.

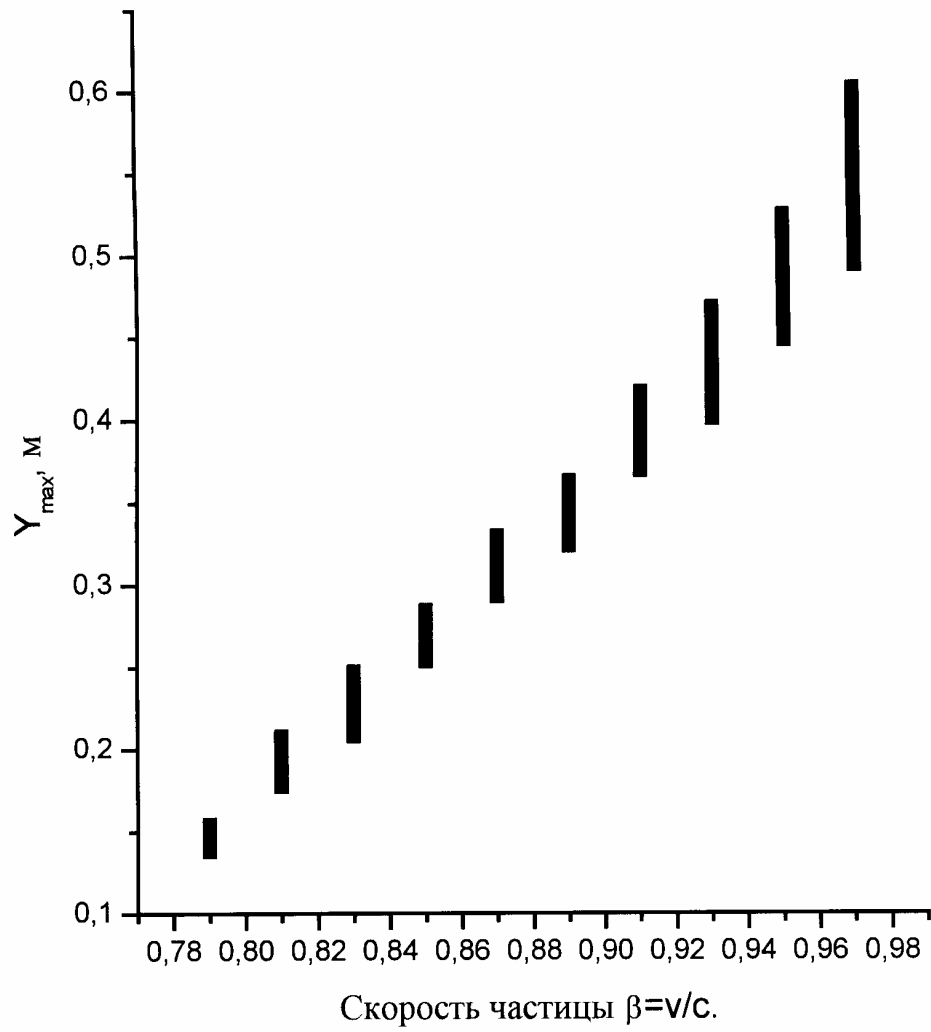


Рисунок 3.

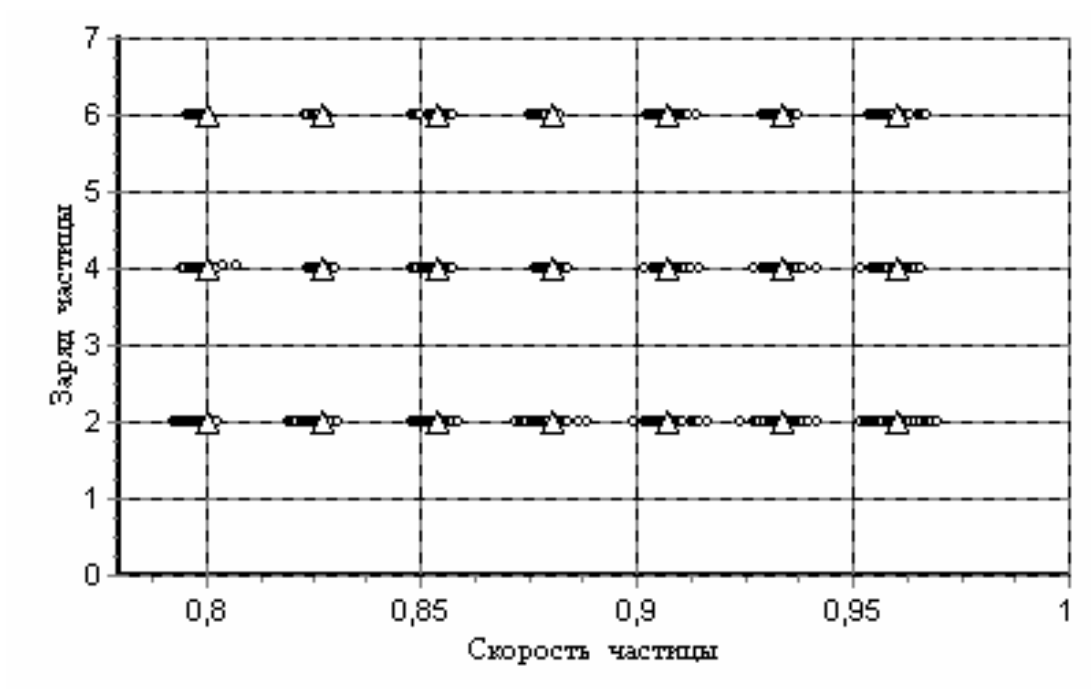


Рисунок 4.

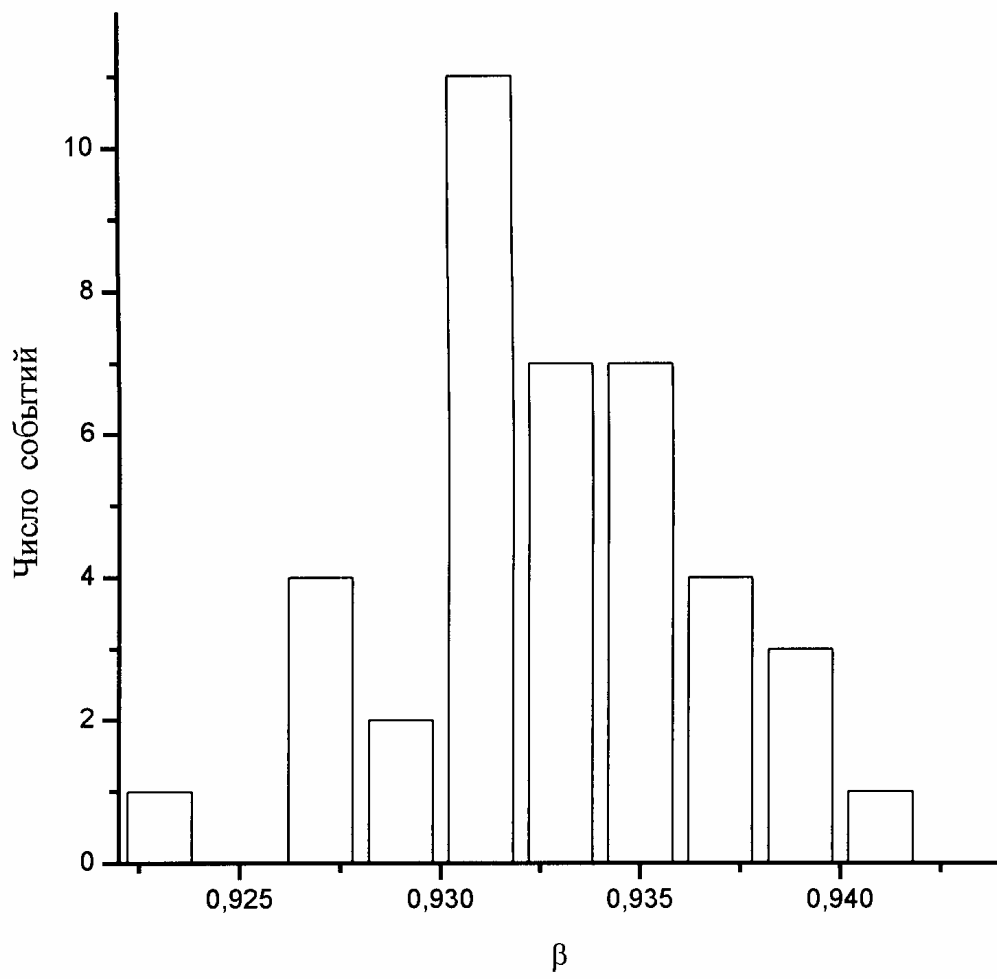


Рисунок 5.