

РАЗРАБОТКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ ФИЗИЧЕСКОГО АНАЛИЗА ДАННЫХ ЭКСПЕРИМЕНТА SELEX

*О. В. Булеков, А. К. Поносов, Д. А. Романов,
А. А. Савченко, В. В. Тарасов, М. А. Файзрахманов*

Московский инженерно-физический институт (государственный университет), Москва

Разработан пакет программного обеспечения для физического анализа результатов эксперимента на адронном спектрометре. Пакет применен для анализа данных эксперимента SELEX.

A software framework for physical analysis of results of experiment on hadron spectrometer has been developed. It is applied on analysis of results of SELEX experiment.

PACS: 29.85.-c

ВВЕДЕНИЕ

Эксперимент SELEX проведен на гиперонном пучке с энергией 600 ГэВ Тэватрона (Лаборатория им. Э. Ферми) для изучения свойств очарованных барионов [1].

Исходное программное обеспечение для эксперимента написано на языке фортран с использованием процедурно-ориентированного подхода. Для физического анализа пользователю предоставляется возможность самостоятельно писать процедуры на фортране, которые используют информацию о событиях, представленную с помощью множества массивов в виде реляционной модели [2].

На C++ на базе пакета ROOT был написан альтернативный программный пакет для физического анализа данных. При разработке пакета были намечены следующие цели:

- увеличение скорости обработки данных;
- оптимизация для использования в сетях GRID [3];
- предоставление пользователю для физического анализа удобного инструментария с объектно-ориентированной архитектурой.

1. ПРОЦЕСС ОБРАБОТКИ ДАННЫХ

На рис. 1 обобщенно представлена структурная схема системы обработки экспериментальных данных и моделирования событий.

Обработка большого количества экспериментальных данных порядка 10^8 – 10^9 триггерных событий занимает значительное время. Обработка одного события существующим программным обеспечением на процессоре с частотой 2400 МГц и производительностью 4800 mips (million instructions per second — миллион инструкций в секунду) требует

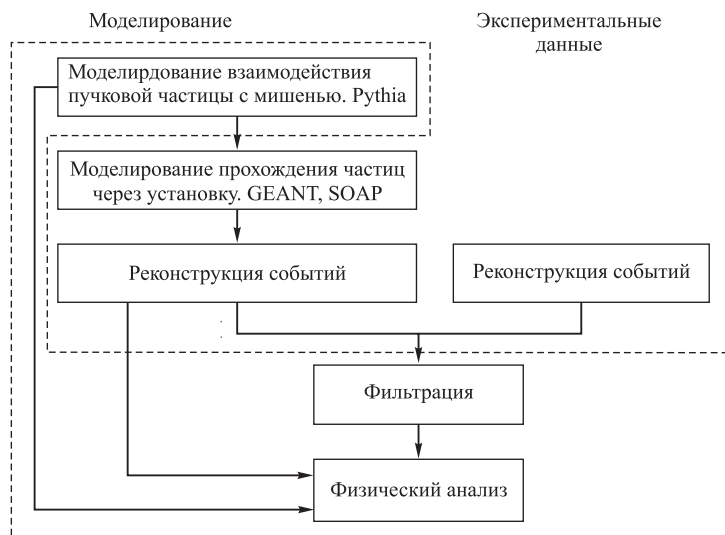


Рис. 1. Структурная схема процесса обработки экспериментальных данных и моделирования событий

около $7 \cdot 10^{-4}$ с процессорного времени. На кластере в ИТЭФ, использующем описанные выше процессоры, полный объем статистики обрабатывается за 7–12 сут, при использовании сети GRID виртуальной организации PHOTON — за 2–5 сут. Поэтому первоначально события отбираются с грубыми критериями отбора, что приводит к сокращению статистики до 10^6 – 10^7 триггерных событий. Для анализа отобранных событий и получения физических результатов используется то же программное обеспечение. Обработка отфильтрованной статистики занимает десятки минут.

При моделировании данных можно выделить следующие основные шаги: моделирование взаимодействия пучковой частицы с мишенью, например, с помощью пакетов PYTHIA или FRITIOF [4]; моделирование прохождения частиц через установку с помощью пакета GEANT [5] или пакета Event Embedding, входящего в состав SOAP (SELEX Off-line Analysis Program); реконструкция событий с помощью пакета SOAP. После этого моделированные данные проходят ту же обработку, что и экспериментальные.

При моделировании наибольшее процессорное время затрачивается на моделирование прохождения частиц через установку.

Штриховой линией на рис. 1 выделена область, объединяющая задачи, для которых может быть использован разработанный пакет. В существующем программном обеспечении параметры рассчитываются независимо от того, проводится ли грубый отбор или конечный физический анализ. После фильтрации с грубыми критериями отбора данные записываются в том же формате, что и исходная статистика. В разработанном пакете многие параметры рассчитываются один раз, а система ROOT-файлов и абстракция классов чтения событий позволяют после проведения грубого отбора записать только данные, необходимые для физического анализа. За счет этого скорость обработки данных в процессе физического анализа в 3–7 раз (в зависимости от заданной логики) выше, чем у исходного программного обеспечения.

2. АРХИТЕКТУРА

В разработанном пакете реализуется архитектура модель–представление–контроллер [6] (рис. 2). В этой архитектуре модель данных, управляющая логика и представление данных пользователю разделены на отдельные компоненты, что облегчает возможность их модификации и использования пакета для различных целей.



Рис. 2. Архитектура модель–представление–контроллер

На рис. 3 представлена упрощенная схема иерархии классов. Пакет включает в себя классы для чтения и записи, преобразования данных в объектную модель [7], классы для отбора событий, классы для представления данных пользователю, интерфейсы для моделирования событий.

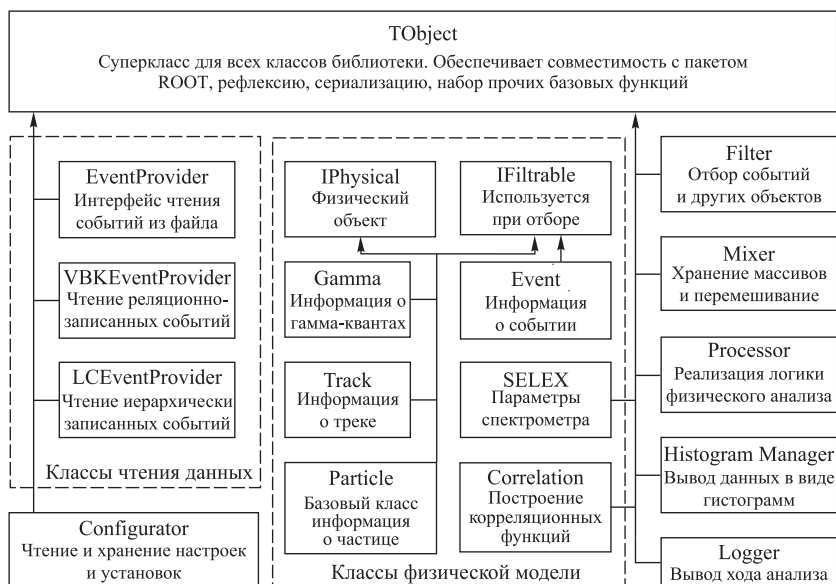


Рис. 3. Упрощенная схема классов

Параметры установки и эксперимента, а также физические константы представлены в виде классов и функций, реализуемых глобальными объектами, что дает им сквозную функциональность для всего приложения.

3. МОДЕЛЬ ДАННЫХ

Наиболее важной частью архитектуры модель–представление–контроллер является модель данных. Она представляется в виде иерархически вложенных объектов, связанных ссылками. На рис. 4 приведена структура объекта данных, представляющая триггерное событие эксперимента. Кроме информации о событии (координаты точки взаимодействия,

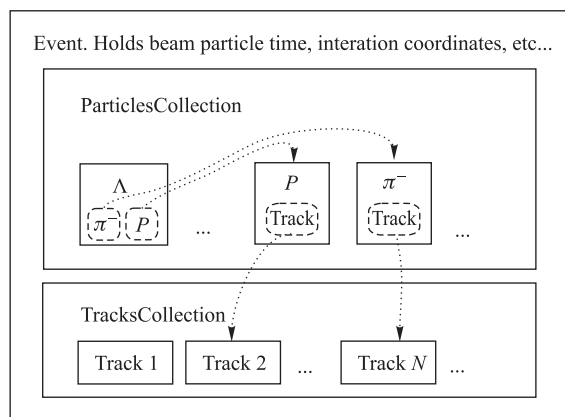


Рис. 4. Объектная модель события эксперимента

тип пучковой частицы и другие), объект включает в себя объекты частиц и треков, которые связаны друг с другом необходимыми ссылками. Например, объекты, представляющие физическую модель частиц, могут иметь ссылку на родительскую частицу, дочерние частицы или на соответствующий трек. Такая модель связи объектов составляется в соответствии с физической моделью события, позволяя эффективно обращаться к данным и удобно программировать логику анализа.

4. ПРИМЕР АЛГОРИТМА АНАЛИЗА

Пользователю предлагается программными методами задать логику физического анализа. В типовом алгоритме анализа точкой входа в приложение является класс, производный от Processor. Для инициализации переменных и задания параметров используется класс Configuration. После инициализации с помощью класса, производного от EventProvider, организуется чтение данных триггерных событий; данные возвращаются пользователю в виде объекта класса Event. Производится отбор необходимых данных, для чего можно использовать классы, производные от Filter. С помощью ряда классов производится обработка и анализ данных, сохранение результатов. Для отображения гистограмм предоставляется класс HistogramManager, позволяющий автоматически строить ряд гистограмм для объектов пакета, хранить гистограммы пользователя, выводить и сохранять все гистограммы. Класс Logger предоставляет доступ к выходному потоку.

Таким образом, с точки зрения архитектуры Модель–представление–контроллер, класс Processor с набором классов, необходимых для получения и обработки данных, выступают в роли контроллера, классы HistogramManager и Logger — в качестве представления, а класс Event, содержащий информацию о событии, частицах и треках, является моделью данных.

На рис. 5 приведен пример алгоритма анализа корреляций пар Λ -гиперонов. Для запуска приложения используется класс, производный от Processor, в котором задана логика анализа. Объект класса EventProvider считывает данные события из файла и возвращает их в виде объекта класса Event. Далее данные события проходят отбор. На

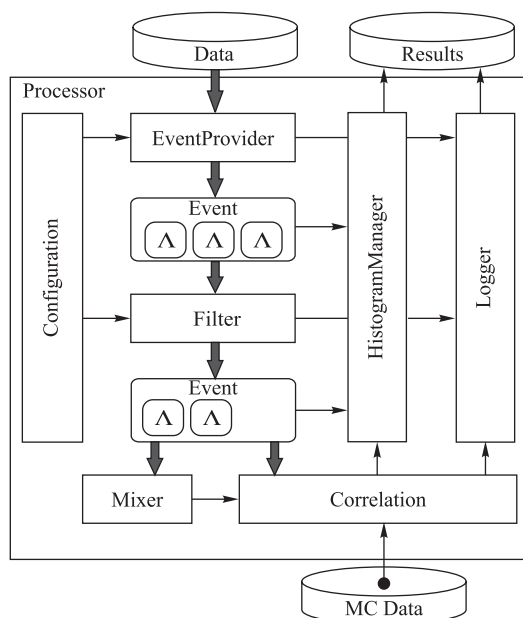


Рис. 5. Пример программы анализа корреляций пар Λ -гиперонов

каждом этапе данные заносятся в гистограммы, ход выполнения программы фиксируется в логах. Частицы, прошедшие отбор, сохраняются в классе *Mixer* и в дальнейшем используются для построения фоновых распределений методом перемешивания. Событие, прошедшее отбор, передается в объект класса *Correlation*, где происходит построение корреляционной функции, для чего используются данные, полученные классом *Mixer*, а также загружаются из файла данные, полученные с помощью моделирования.

5. РЕГИСТРАЦИЯ ПАР Λ^0

На основе статистического материала эксперимента SELEX представлены результаты предварительного отбора пар Λ -гиперонов.

Были отобраны события, содержащие не менее двух Λ -гиперонов, идентифицированных по распаду на $p\pi^-$.

Отбирались распады в вершинном детекторе.

Требовалось, чтобы минимальное расстояние между треками V^0 составляло менее 0,1 см.

Эффективная масса системы $p\pi^-$ должна была лежать в пределах $1,10 < M_{\text{эфф}} < 1,13 \text{ ГэВ}/c^2$.

Угол ξ , образованный вектором импульса Λ -гиперона и прямой, проходящей через вершину рождения и вершину распада Λ -гиперона (рис. 6), должен был быть менее 0,1 мрад.

Параметры распределения Подольянского–Арментероса [8] должны были соответствовать возможным значениям для Λ -гиперонов.

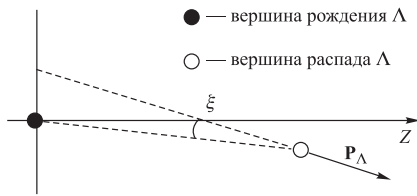


Рис. 6. Угол ξ

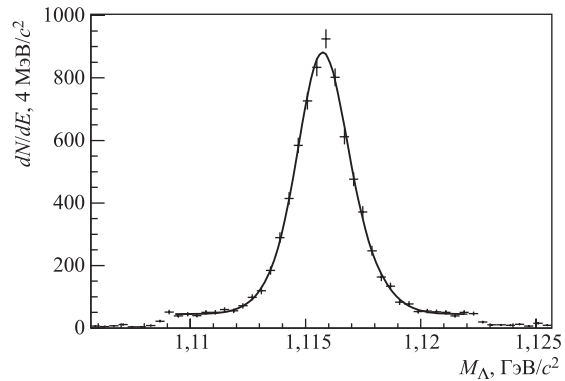


Рис. 7. Спектр эффективных масс $p\pi^-$. Отобрано 12788 пар. $M_\Lambda = (1115,70 \pm 0,06) \text{ МэВ}/c^2$, $\sigma = (0,91 \pm 0,15) \text{ МэВ}/c^2$

В результате анализа $8 \cdot 10^8$ триггерных событий, полученных с помощью очарованного триггера, количество пар Λ -гиперонов, прошедших предварительный отбор, составило 12788.

На рис. 7 представлено распределение по эффективной массе системы $p\pi^-$. Распределение фитировалось суммой функции Гаусса и полинома второй степени для описания фона. Полученное значение $M_\Lambda = (1115,70 \pm 0,06) \text{ МэВ}/c^2$, $\sigma = (0,91 \pm 0,15) \text{ МэВ}/c^2$.

На рис. 8 для пар $\Lambda\Lambda$ представлена зависимость эффективной массы системы $p\pi^-$, соответствующей первому Λ -гиперону, от эффективной массы системы $p\pi^-$, соответствующей второму Λ -гиперону. Фон составляет менее 6% от сигнала.

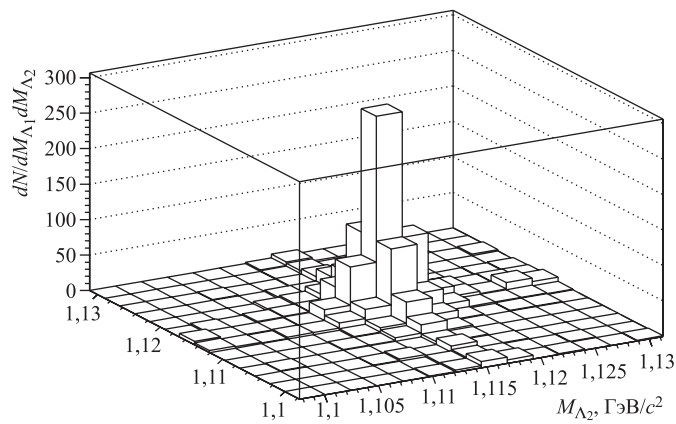


Рис. 8. Двумерное распределение зависимости эффективной массы системы $p\pi^-$, соответствующей первому Λ -гиперону, от эффективной массы системы $p\pi^-$, соответствующей второму Λ -гиперону

ЗАКЛЮЧЕНИЕ

Разработан пакет программного обеспечения для физического анализа результатов экспериментов на адронных спектрометрах. Пакет применен для анализа данных эксперимента SELEX.

Пакет рассчитан на обработку данных большого объема (10^6 – 10^9 событий) и оптимизирован для работы в сетях GRID. Скорость обработки данных в 3–7 раз (в зависимости от заданной логики) выше, чем у исходного программного обеспечения.

На основе данных эксперимента SELEX была получена рекордная мировая статистика пар Λ -гиперонов, составляющая 12788 событий.

Авторы статьи благодарят сотрудников ИТЭФ А.Г. Долголенко, В.А. Матвеева и Г.В. Давиденко за помощь в работе, а также коллаборацию SELEX за предоставленную возможность анализа событий.

Работа выполнена при поддержке гранта РФФИ (проект № 08-02-01014а).

СПИСОК ЛИТЕРАТУРЫ

1. *Balatz M. Y. et al. (SELEX). The Lead-Glass Calorimeter for the SELEX Experiment. FERMILAB-TM-2252. 2004. 42 p.*
2. *Godd E. F. A Relational Model of Data for Large Shared Data Banks // Commun. of the ACM 13. 1970. No. 6. P. 377–387.*
3. *Gagliardi F. et al. Building an Infrastructure for Scientific Grid Computing: Status and Goals of the EGEE Project // Philos. Trans.: Math., Phys. Engin. Sci. 2005. No. 363. P. 1729–1742.*
4. *SjOstrand T. PYTHIA 5.7 and JETSET 7.4 (Physics Manual). CERN-TH.7112. 1993.*
5. *Brun R., Bruyant F., Marie M. GEANT Detector Description and Simulation Tool. Application Software Group, CND. CERN Program Library Long Writeup W5013. 1994.*
6. *Fowler, Martin. Patterns of Enterprise Application Architecture. Addison-Wesley Professional, 2002.*
7. *Грэхем И. Объектно-ориентированные методы. Принципы и практика. М.: Вильямс, 2004. 880 с.*
8. *Podolansky J., Armenteros R. // Philos. Mag. 1954. No. 13.*

Получено 17 марта 2008 г.