

Сравнение подходов к разрешению неоднозначностей при реконструкции треков заряженных частиц

А. С. Камкин^{1,2}, П. Н. Белецкий^{1,2*}, В. В. Бурдельная^{1,2}, И. А. Козмин^{1,2}, Д. А. Михалевич^{1,2},
Д. М. Эркенова^{1,2}

¹ ФГБОУ ВО «Российский экономический университет имени Г.В. Плеханова, г. Москва, Российская Федерация

Адрес: 115054, Российская Федерация, г. Москва, Стремянный пер., д. 36

² ФГБУН Институт системного программирования им. В.П. Иванникова РАН, г. Москва, Российская Федерация

Адрес: 109004, Российская Федерация, г. Москва, ул. А. Солженицына, д. 25

* belecky@ispras.ru

Аннотация

При решении задачи реконструкции треков заряженных частиц возможны неопределённости — ситуации, когда нельзя однозначно разделить множество точек на треки (выяснить, к какому треку принадлежит та или иная точка, или установить, относятся ли распознанные треки к одному реальному треку). Обработка таких ситуаций осуществляется в конце реконструкции треков, после того как построено множество треков-кандидатов (с помощью фильтра Калмана или иного метода). Основные цели: (1) выявить и удалить ложные треки (треки-кандидаты, не имеющие реальных прототипов); (2) устранить избыточность (минимизировать число треков-кандидатов, соответствующих одному реальному треку). Последняя цель достигается путём кластеризации треков-кандидатов и выбора лучшего представителя из каждого кластера (в более общем случае представитель конструируется путем «соединения» треков кластера). В статье рассматриваются некоторые подходы к разрешению неоднозначностей (как алгоритмические, так и использующие методы машинного обучения); приводится их сравнение по таким показателям, как эффективность, доля ложных треков, доля треков-дубликатов. Анализ проводится в среде MPDRoot, разрабатываемой для эксперимента MPD проекта NICA, однако предложенные подходы являются универсальными и могут применяться в других экспериментах.

Ключевые слова: сверточная нейронная сеть, оптимизация гиперпараметров, Optuna, NupurOpt

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Для цитирования: Сравнение подходов к разрешению неоднозначностей при реконструкции треков заряженных частиц / А. С. Камкин [и др.] // Современные информационные технологии и ИТ-образование. 2024. Т. 20, № 3. С. 666-677. <https://doi.org/10.25559/SITITO.020.202403.666-677>

© Камкин А. С., Белецкий П. Н., Бурдельная В. В., Козмин И. А.,
Михалевич Д. А., Эркенова Д. М., 2024



Контент доступен под лицензией Creative Commons Attribution 4.0 License.
The content is available under Creative Commons Attribution 4.0 License.



Comparison of Ambiguity Resolution Approaches for Reconstructing Charged Particle Tracks

A. S. Kamkin^{a,b}, P. N. Beletsky^{a,b*}, V. V. Burdelnaya^{a,b}, I. A. Kozmin^{a,b}, D. A. Mikhalevich^{a,b}, J. M. Erkenova^{a,b}

^a Plekhanov Russian University of Economics, Moscow, Russian Federation
Address: 36 Stremyanny lane, Moscow 115054, Russian Federation

^b Ivannikov Institute for System Programming of the Russian Academy of Sciences, Moscow, Russian Federation

Address: 25 Alexander Solzhenitsyn St., Moscow 109004, Russian Federation

* belecky@ispras.ru

Abstract

The key task of most high-energy physics experiments is to reconstruct the trajectories of charged particles, the so-called tracks. The first stage of event reconstruction is the search for candidate tracks, often performed by applying the Kalman filter. Candidate tracks often contain duplicate tracks and fake tracks, which complicates further interpretation of the experimental results. That is why the next step in the reconstruction of a high-energy physics event is ambiguity resolution. The main objective of this paper is to eliminate data redundancy and reduce the number of fake tracks, that is, tracks that do not correspond to real particles. We describe the subject area of the study, outline the structure of the TPC detector of the MPD experiment, and analyze existing methods for resolving track ambiguities. The scientific novelty of this paper is the development of new algorithms for ambiguity resolution. The analysis is performed in the MPDRoot environment developed for the MPD experiment of the NICA project, but the proposed approaches are universal and can be applied in other experiments. To test our algorithms, we simulate 10,000 events of the MPD experiment by Monte Carlo simulation and apply the developed ambiguity resolution methods on the simulation data. The ambiguity resolution methods are compared in terms of efficiency, fake track rate, and duplicate track rate. Since experiment-specific test data were used, the conducted research allows us to determine the most appropriate ambiguity resolution method for the MPD experiment.

Keywords: high energy physics, particle detector, event reconstruction, track recognition, ambiguity resolution, NICA project, MPD experiment

Conflict of interests: The authors declares no conflict of interest.

For citation: Kamkin A.S., Beletsky P.N., Burdelnaya V.V., Kozmin I.A., Mikhalevich D.A., Erkenova J.M. Comparison of Ambiguity Resolution Approaches for Reconstructing Charged Particle Tracks. *Modern Information Technologies and IT-Education*. 2024;20(3):666-677. <https://doi.org/10.25559/SITITO.020.202403.666-677>

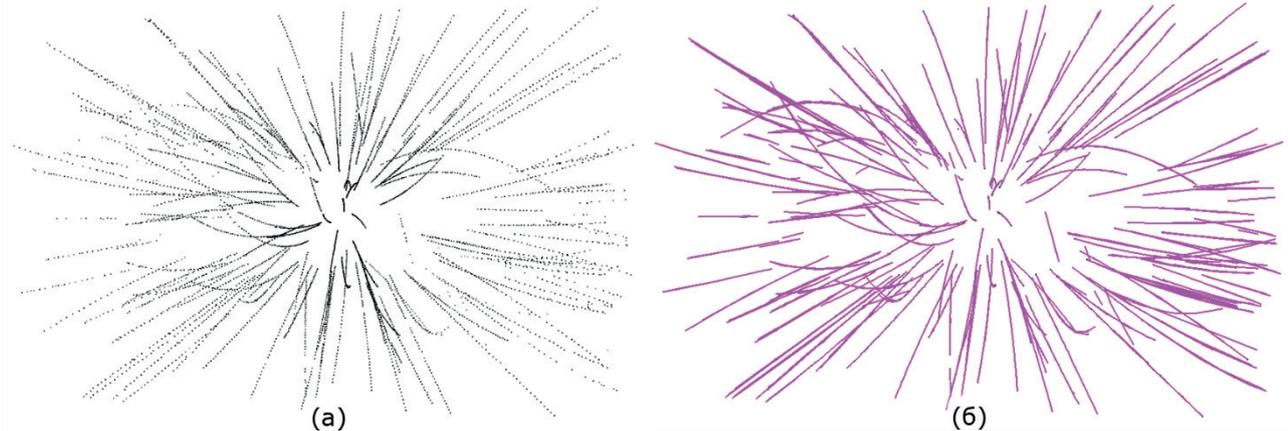


Введение

Событие в экспериментальной физике высоких энергий – это процесс рождения вторичных частиц в результате столкновения первичных ускоренных частиц, зарегистрированный с помощью детекторов. Для получения информации о рожденных частицах используются математические методы для нахо-

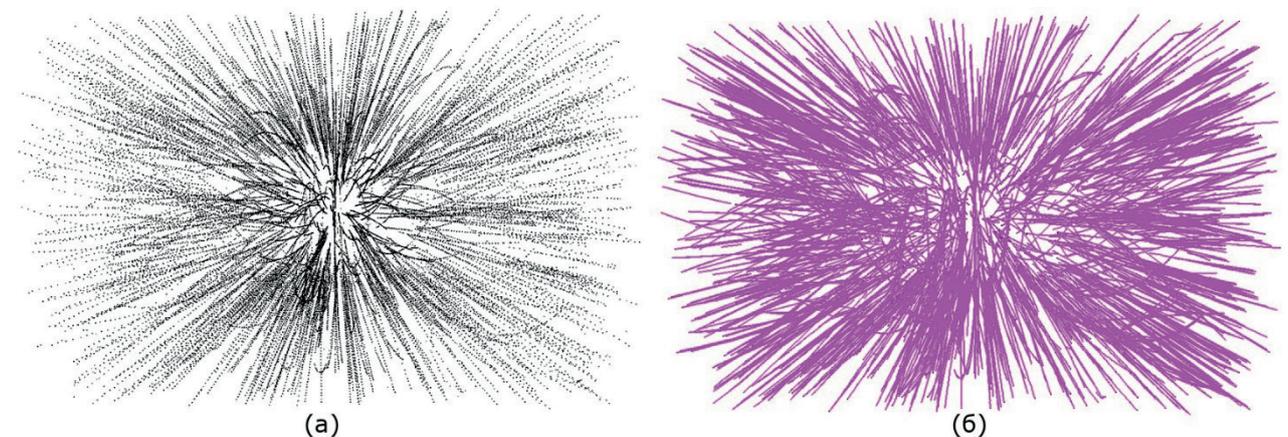
ждения и реконструкции траекторий движения заряженных частиц, называемых треками, а также специальные трековые детекторы¹.

Задача реконструкции треков частиц заключается в сопоставлении отдельных откликов детектора (хитов) трекам различных частиц [1]. Примеры хитов и соответствующих им реконструированных треков приведены на рисунках 1 и 2.



Р и с. 1. Событие периферического взаимодействия Au+Au при энергии столкновения 7 ГэВ/нуклон (~100 треков): (а) хиты, (б) найденные треки
F i g. 1. Au+Au peripheral interaction event at 7 GeV/nucleon collision energy (~100 tracks): (a) hits, (b) found tracks

Источник: здесь и далее в статье все таблицы и рисунки составлены авторами.
Source: Hereinafter in this article all tables and figures were made by the authors.



Р и с. 2. Событие центрального взаимодействия Au+Au при энергии столкновения 7 ГэВ/нуклон (~1000 треков): (а) хиты, (б) найденные треки
F i g. 2. Au+Au central interaction event at 7 GeV/nucleon collision energy (~1000 tracks): (a) hits, (b) found tracks

Среди найденных в ходе реконструкции треков могут быть треки-дубликаты и ложные треки, составленные из хитов других треков. Для очистки данных, полученных при реконструкции треков, выполняется процедура разрешения неоднозначностей (*ambiguity resolution*). Она позволяет снизить избыточность данных, сохранив информацию о реальных тре-

ках, а также минимизировать число ложных треков [2]. Несмотря на то, что анализ методов, который проводится в данной работе, не привязан к конкретному эксперименту, тестовые данные для анализа специфичны для эксперимента MPD [3]. В связи с этим, проведённый анализ позволяет выбрать наиболее подходящий метод разрешения неоднознач-

¹ Gessinger-Befurt P. Development and improvement of track reconstruction software and search for disappearing tracks with the ATLAS experiment : Dissertation zur Erlangung des Grades Doktor der Naturwissenschaft. Mainz, JGU-PUBLICATIONS, 2021. 325 p. <http://doi.org/10.25358/openscience-5901>



ностей для данного эксперимента.

Научная новизна статьи состоит в разработке новых алгоритмов разрешения неоднозначностей.

В статье приводятся результаты сравнения методов по следующим критериям: эффективности нахождения реальных треков эффективности (*efficiency*), доли треков-дубликатов (*duplicate rate*) и доли ложных треков (*fake rate*).

Оставшаяся часть статьи организована следующим образом. В разделе 1 определяются основные понятия, определяются показатели оценки качества реконструкции треков. Раздел 2 посвящён анализу существующих методов разрешения неоднозначностей: описывается общая схема и приводится описание алгоритмов разрешения неоднозначностей различных экспериментов. В разделе 3 описываются предлагаемые подходы и осуществляется классификация рассмотренных алгоритмов. В разделе 4 проводится сравнительный анализ методов: описывается методология исследования и получение тестовых данных, формулируется критерий сравнения, приводятся результаты сравнения. Раздел 5 является заключительным, в нём описываются результаты исследования: выявляется лучший согласно критерию сравнения метод, приводятся значения показателей оценки качества восстановления треков, а также задаётся направление дальнейших исследований.

1. Предметная область

1.1 Реконструкция события

Реконструкция треков состоит из нескольких этапов:

- поиск треков-кандидатов;
- фитирование, или оценка параметров трека;
- разрешение неоднозначностей.

Выделяют *глобальные* и *локальные методы* поиска треков-кандидатов [4]. В глобальных методах все хиты обрабатываются как единая совокупность, и результат реконструкции не зависит от того, какой хит обрабатывался первым и общей последовательности обработки остальных хитов. К глобальным методам относятся метод поиска шаблонов (*template matching*), преобразование Хафа, метод искусственной сетчатки (*artificial retina*), а также класс методов реконструкции треков, основанных на нейронных сетях [5-9]. В локальных методах реконструкция треков-кандидатов начинается с *поиска первичных трековых сегментов (seeding)*, то есть на начальном этапе определяются комбинации хитов, которые порождены одной и той же заряженной частицей [10]. *Первичный трековый сегмент (seed)* представляет собой часть трека, которая состоит из небольшого количества хитов (обычно двух или трёх) [11]. Первичные трековые сегменты могут генерироваться во внутренней части трекового детектора, где измерения часто имеют очень высокую точность, или во внешней части, где плотность треков ниже [12]. В отличие от глобальных методов, результат работы таких алгоритмов зависит от того, с каких хитов начался процесс реконструкции трека и в каком порядке обрабатывались остальные хиты. Наиболее распространённым локальным методом поиска треков и оценки их параметров является фильтр Калмана [13]. На основании параметров первичного трекового сегмента делается предсказание следующего хита, который мог бы быть отнесён к данному треку-кандидату. Если удаётся найти хит, который

соответствует предсказанию, он включается в треки-кандидаты. Параметры трека-кандидата обновляются с использованием информации добавленного хита. Процедура является итеративной [14].

Фитирование необходимо для определения кинематических параметров частицы, сопоставляемой треку. Для оценки параметров частицы обычно применяется метод наименьших квадратов [15].

Из-за комбинаторного характера поиска множество треков-кандидатов часто бывает большим и содержит *ложные треки (fake tracks)* и *треки-дубликаты (duplicate tracks)* [16]. Считается, что треки-кандидаты *соответствуют реальному треку (truth-matched track-candidate)*, если в нём не менее некоторой доли хитов α (обычно 50%) принадлежит реальному треку. В противном случае трек называется ложным [17]. Бывают ситуации, когда одной реальной частице соответствует несколько треков-кандидатов – тогда говорят о треках-дубликатах. Наличие ложных треков и треков-дубликатов затрудняет анализ события физики высоких энергий. Поэтому важным этапом в реконструкции треков частиц является этап разрешения неоднозначностей, при котором выявляются и удаляются ложные треки, а также устраняется избыточность.

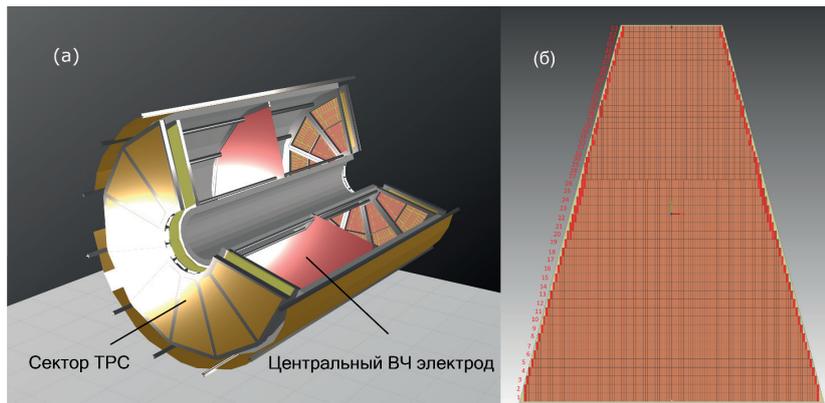
1.2 Детектор ТРС эксперимента MPD

MPD (*Multi-Purpose Detector*) – это эксперимент по столкновению тяжелых ионов на ионном коллайдере NICA, один из флагманских проектов, который планируется ввести в эксплуатацию в Объединенном институте ядерных исследований в Дубне [18]. Его целью является поиск новых физических явлений при столкновении тяжелых ядер в диапазоне энергий $4 \leq \sqrt{s_{NN}} \leq 11$ ГэВ [19]. В данной работе исследуется реконструкция треков в ТРС. Детектор ТРС (*Time Projection Chamber*) является основным трекингвым детектором в центральной области быстрой эксперимента MPD. ТРС представляет собой полный цилиндр, заполненный газовой смесью, в торцах которого расположены многопроволочные камеры для регистрации данных. Частицы, проходящие через объем ТРС, оставляют облака заряженных электронов и ионов, которые дрейфуют под влиянием электрического поля к торцам камеры, где детектируются плоскостью с чувствительными элементами (пэдами). Длина ТРС составляет 3.4 метра, рабочая область находится в пределах от минимального радиуса 0.27 метра до максимального радиуса 1.4 метра. Каждый из торцов состоит из 12 секторов с пэдами. Сектор состоит из 53 рядов пэдов (см. рис. 3). ТРС рассчитан на детектирование заряженных частиц в диапазоне псевдобыстрот $|\eta| < 1.2$. Он будет обеспечивать разрешение по импульсу около 3% для заряженных частиц с поперечным импульсом в диапазоне $0.1 < p_t < 1$ ГэВ/с [19].

1.3 Оценка эффективности реконструкции

Для оценки эффективности реконструкции события обычно проводится моделирование столкновения элементарных частиц методом Монте-Карло. Такое моделирование позволяет получить множество хитов, для которых известно, к каким частицам они принадлежат. Треки, полученные в рамках моделирования, будем называть реальными. Характеристики качества реконструкции рассчитываются по первичным реальным трекам, удовлетворяющим следующим условиям:





Р и с. 3. MPD TPC: (а) схематический вид TPC, (б) сектор TPC

F i g. 3. MPD TPC: (a) schematic view of TPC, (b) TPC sector

- $|\eta| < 1.5$;
- $p_t > 100$ МэВ/с;
- число хитов в треке ≥ 9 .

Хиты используются для реконструкции треков, включая этап разрешения неоднозначностей. Треки, полученные после реконструкции, называются треками-кандидатами. Для расчета характеристик качества реконструкции треков используются только треки-кандидаты, число хитов в которых больше либо равно заданной константе (в данной работе принимается 6 хитов). Сопоставление реальных треков и треков-кандидатов позволяет вычислить предложенные для оценки работоспособности алгоритма разрешения неоднозначностей показатели – эффективность, долю ложных треков и долю треков-дубликатов.

Определим распознанный трек как реальный трек, которому соответствует хотя бы один трек-кандидат. При этом процентное соотношение числа хитов трека-кандидата, относящихся к реальному треку, должно быть больше либо равно заданной константе (в данной работе 50%).

Эффективность реконструкции треков (*efficiency*) рассчитывается как процентное соотношение числа распознанных треков к числу реальных треков.

Доля ложных треков рассчитывается как процентное соотношение числа ложных треков-кандидатов к числу треков-кандидатов.

Доля дубликатов рассчитывается как процентное соотношение числа треков-дубликатов к числу треков-кандидатов.

2. Методы разрешения неоднозначностей

2.1 Общая схема

Все рассматриваемые алгоритмы разрешения неоднозначностей реконструкции треков заряженных частиц так или иначе основываются на кластеризации треков и на последующем выборе лучшего или лучших треков в кластере. В более общем случае, представитель конструируется путем «соединения» треков кластера.

2.2 Существующие подходы

Существующие подходы можно условно разделить на методы с динамической и предварительной кластеризацией. К методам с динамической кластеризацией относится, например, Greedy Solver [20]. К методам с предварительной кластеризацией относятся методы разрешения неоднозначностей в эксперименте ATLAS, в которых в том числе могут применяться модели машинного обучения [21].

2.2.1 Алгоритм Greedy Solver

Greedy Solver относится к алгоритмам с динамической кластеризацией [20]. Для алгоритма задается максимальное число общих хитов, которое трек может иметь с другими треками. Под общими хитами понимаются хиты трека-кандидата, которые встречаются в других треках-кандидатах. Первым шагом в алгоритме отбираются все треки, имеющие больше общих хитов, чем было предварительно задано. Далее среди отобранных треков удаляется трек с наибольшим отношением числа общих хитов к числу хитов в треке. В случае, если у двух треков отношение одинаковое, удаляется трек с наибольшим χ^2 . Следующим шагом для всех треков пересчитывается число общих хитов. Процедура повторяется до тех пор, пока число общих хитов оставшихся треков не будет меньше или равным изначально заданному параметру.

2.2.2 Метод Pattern Recognition

Алгоритм Pattern Recognition является одним из методов динамической кластеризации². На первом этапе его реализации все треки сортируются по числу хитов. Начиная с самого длинного в хитах трека-кандидата все хиты трека помечаются как «использованные». Для каждого последующего трека проверяется, как много хитов, принадлежащих треку, были ранее отмечены как «использованные». Если доля таких хитов в рассматриваемом треке превышает заданное значение (в работе оно равно 30%), то трек-кандидат исключается. Если трек не исключен, то все его хиты отмечаются как «использованные». Процедура повторяется, пока не будут рассмотрены все треки-кандидаты. В разделе 3 приводится аналогичный метод, который был разработан независимо от приведенного выше.

² Gorbounov I. Pattern Recognition in the Inner Tracking System of HERA-B and Measurement of the V_p Production Cross Section in pN Collisions : Dissertation zur Erlangung des Grades eines Doktors der Naturwissenschaften. Siegen: Universitat Siegen, 2003. 108 p. URL: <https://inspirehep.net/literature/672761> (дата обращения: 28.08.2024).



2.2.3 Метод Neural-Network Pixel Clustering

В методе разрешения неоднозначностей треков проекта ACTS используются алгоритмы машинного обучения³. Треки сортируются по убыванию числа хитов, а затем разбиваются на кластеры. Кластер формируется из рассматриваемого трека, а также других треков, которые имеют общие хиты с рассматриваемым. Оценка треков определяется предсказаниями нейронной сети. На вход нейросети принимаются параметры трека:

- число хитов, формирующих трек-кандидат (*measurements*);
- число выбросов в треке (*outliers*);
- уменьшенная статистика χ^2 (χ^2/NDF);
- поперечный угол (ϕ);
- псевдобыстрота (η);
- число состояний (*number of states*);
- число «дырок» (*number of holes*);
- число степеней свободы трека (*NDF*).

Числом недостающих измерений («дырок») трека называется количество предсказаний фильтра Калмана, которым не был сопоставлен хит. Выбросом называется такой хит, который при добавлении в трек-кандидат повышает статистику χ^2 трека на большее значение, чем установлено в качестве допустимого (в данной работе допустимое значение было зафиксировано равным 30). Сумма хитов трека-кандидата, числа недостающих измерений и числа выбросов формирует значение числа состояний трека. Число степеней свободы трека равно удвоенной сумме числа хитов трека и числа выбросов в треке. Также добавляется бинарная разметка: трек либо «хороший», либо нет. Если одному реальному треку соответствует два и более треков-кандидатов, то «хорошим» будет называться тот трек-кандидат, у которого наибольшее число хитов, соответствующих реальному треку, наименьшее число выбросов среди всех треков-кандидатов и самый маленький χ^2 . Для каждого кластера выбирается трек с наибольшей оценкой, остальные треки исключаются из дальнейшего рассмотрения.

2.2.4 Метод Neural Network Track Clustering

В подходе, предложенном для эксперимента ATLAS, предлагается обучать нейронную сеть для последующей классификации треков⁴. Данными для нейросети служит информация о 838 792 треках-кандидатах: 75% записей используются в качестве обучающей выборки, а оставшиеся 25% – в качестве тестовой. На вход нейросети принимаются параметры трека: число хитов, формирующих трек-кандидат, число выбросов в треке, число хитов, встречающихся в других треках-кандидатах, уменьшенная статистика χ^2 , поперечный импульс, псевдобыстрота, а также три бинарных метки – является ли трек «хорошим» (то есть не ложным и не треком-дубликатом), является ли трек ложным, является ли трек дубликатом.

Результатом работы нейросети являются вероятности принадлежности трека классам «хороший трек», «ложный трек», «трек-дубликат». Трек классифицируется как класс с наибольшей вероятностью.

Перспективным авторы описанного метода считают также использование *двунаправленной сети долго-краткосрочной памяти (Bi-directed Long short-term memory, Bi-directed LSTM)*. Двунаправленные LSTM имеют слой специальных блоков памяти, благодаря чему они способны запоминать информацию на долгие периоды времени [22]. Входной слой нейросети последовательно принимает описанные ранее характеристики треков-кандидатов. В работе предлагается передавать во входной слой характеристики треков, отсортированные по значению псевдобыстроты (η). Результатом работы LSTM является многоклассовая классификация треков.

2.2.5 Алгоритм в проекте MPDRoot

В проекте MPDRoot используется следующий метод разрешения неоднозначностей треков [23]. В результате работы алгоритма поиска треков для каждого трека-кандидата известно значение χ^2 . Треки-кандидаты сортируются по убыванию числа хитов. Далее происходит попарный перебор треков-кандидатов t_{c1} , t_{c2} . Треки t_{c1} и t_{c2} считаются дубликатами если число их общих хитов не меньше, чем

$$\left\lceil \frac{\min(L(t_{c1}), L(t_{c2}))}{2} \right\rceil,$$

где $\lceil \cdot \rceil$ – округление к большему. Если t_{c1} и t_{c2} являются дубликатами, то удаляется трек, состоящий из меньшего количества хитов. Если t_{c1} и t_{c2} состоят из одинакового количества хитов, то удаляется трек с максимальным χ^2 ⁵.

3. Предлагаемые подходы

В рамках данной статьи были разработаны или улучшены шесть различных методов разрешения неоднозначностей.

3.1 Алгоритм разрешения неоднозначности путем покрытия хитов

В алгоритме разрешения неоднозначности путем покрытия хитов (*Hit-coverage-based track filtering*, далее HCF) используются два параметра:

- отношение числа неиспользованных хитов трека к числу хитов в треке (*newHitsRatio*);
- число идущих подряд неиспользованных хитов (*newHitsInRow*).

Треки-кандидаты сортируются по числу хитов по убыванию. Далее происходит перебор треков в заданном выше порядке и все хиты трека помечаются как «использованные». Для каждого последующего трека вычисляется отношение неис-

³ Ambiguity resolution algorithm based on machine learning [Электронный ресурс] // ACTS. CERN, 2024. URL: <https://acts.readthedocs.io/en/latest/plugins/MLAlgorithms.html#ambiguityresolutionmlalgorithm> (дата обращения: 28.08.2024); Allaire C., Grasland H., Rousseau D., Bouvet F. Machine Learning for Ambiguity Resolution in Acts // Report at the 26th international conference on computing in high energy & nuclear physics (CHEP2023). Norfolk, Virginia, USA, 2023. Available at: <https://indico.jlab.org/event/459/contributions/11453/attachments/9329/13530/Machine-learning-for-Ambiguity-Resolution-in-Acts.pdf> (дата обращения: 28.08.2024).

⁴ Nupur O. Deep Learning for Ambiguity Resolution in ACTS [Электронный ресурс] // IRIS-HEP Fellowship Summer 2021. CERN, 2021. Available at: <https://indico.cern.ch/event/1071399/contributions/4505116/attachments/2306008/3923111/Deep%20Learning%20for%20Ambiguity%20Resolution%20in%20ACTS.pdf> (дата обращения: 28.08.2024).

⁵ Реализация алгоритма удаления треков-дубликатов в MPDRoot [Электронный ресурс] // GitLab, 2024. URL: <https://git.jinr.ru/nica/mpdroot/blob/7f311b2936e5d00f21a20890c19467f017e01367/reconstruction/tracking/lheTrack/MpdTrckKalmanFilter.cxx#L1963> (дата обращения: 28.08.2024).



пользованных хитов к числу хитов рассматриваемого трека, а также подсчитывается число сегментов трека, состоящих из *newHitsInRow* идущих подряд неиспользованных хитов. Затем, если вычисленное отношение меньше заданного *newHitsRatio* или не удалось найти ни одного сегмента трека, состоящего из подходящей количества хитов с подряд идущими неиспользованными хитами, данный трек исключается из дальнейшего рассмотрения. В противном случае все хиты трека помечаются «использованными».

3.1.1 Поиск оптимальных параметров для HCF

Для данного алгоритма был произведен поиск оптимальных параметров методом перебора по сетке в диапазоне:

- для *newHitsRatio*: от 0 до 1 с шагом 0.01;
- для *newHitsInRow*: от 1 до 100 с шагом 1.

Наибольшая эффективность реконструкции треков достигается при значении параметра *newHitsRatio*, равном 0.25, и при *newHitsInRow*, равном 6.

3.2 Алгоритмы попарного сопоставления треков

В алгоритмах попарного сопоставления треки попарно сравниваются друг с другом. Для каждой пары подсчитывается число хитов, присутствующих одновременно в двух треках. Затем вычисляется отношение количества общих хитов к количеству хитов наименьшего трека пары, далее это значение сравнивается с неким заданным параметром *similarityFactor* (в работе равен 0.5). Если найденное отношение превышает зафиксированный *similarityFactor*, пара треков содержит неоднозначность.

3.2.1 Алгоритм попарного сопоставления и отбора треков

Алгоритм попарного сопоставления и отбора треков или *Pairwise track matching and merging* (далее PWS) разрешает неоднозначность за счёт выбора лучшего трека из пары. Какой из треков считается лучшим определяется по количеству хитов в каждом из треков. Тот, в котором их больше считается лучшим. Если количество хитов равно, то выбирается случайный трек.

3.2.2 Алгоритм попарного сопоставления и слияния треков

Данный алгоритм (*Pairwise track matching and merging*, далее PWM) имеет иной подход относительно алгоритма 3.2.1 в разрешении неоднозначностей. После обнаружения неоднозначности в паре треков, PWM выбирает все уникальные хиты исходных двух треков и формирует из них новый трек. Два исходных трека удаляются.

3.3 Алгоритмы на основе графов

После того, как предложенные алгоритмы попарного перебора 3.2 показали хорошие результаты эффективности реконструкции, были разработаны улучшенные по времени работы алгоритмы, основанные на идее представления трека в виде графа близости. Граф хранит информацию о том, скольким трекам принадлежит каждый хит. Это является первым шагом для всех алгоритмов данной секции.

3.3.1 Алгоритм отбора треков путем построения графа близости

Данный алгоритм (*Track proximity graph selection*, далее PGS) совершает обход в глубину по полученному графу близости и получает все возможные множества пересекающихся треков. Из каждого полученного множества выбирается один лучший

трек исходя из критерия количества хитов в треке, где больше хитов – лучше.

3.3.2 Алгоритм слияния треков путем построения графа близости

Данный алгоритм (*Track proximity graph merging*, далее PGM) является вариацией алгоритма 3.3.1. Единственное отличие PGM заключается в работе с множеством пересекающихся треков. На последней стадии PGM не выбирает лучший трек из имеющихся, а формирует его из всех уникальных хитов пересекающихся треков.

3.4 Алгоритм разрешения неоднозначности с применением машинного обучения

В ходе исследования существующих подходов было найдено несколько методов с использованием нейронных сетей для разрешения неоднозначностей треков. За основу алгоритма разрешения неоднозначности с применением машинного обучения берется метод *Neural Network Pixel Clustering 2.2.3*. Поскольку в оригинальной статье отсутствуют детали реализации, была предложена собственная реализация метода, получившая название *Neural network-based selection* (далее NNS) [25]. Ключевым отличием NNS от первоисточника является переработанная функция потерь.

$$L(y_{\text{true}}, y_{\text{pred}}) = \frac{1}{N} \sum_{i=1}^N \max(0, y_{\text{true}}^{(i)} - y_{\text{pred}}^{(i)} + \text{margin}),$$

где:

- $L(y_{\text{true}}, y_{\text{pred}})$ – значение функции потерь;
- N – общее число образцов в выборке;
- $y_{\text{true}}^{(i)}$ – истинное значение целевой переменной для i -го образца;
- $y_{\text{pred}}^{(i)}$ – предсказанное моделью значение целевой переменной для i -го образца;
- margin – параметр смещения, равный 0.05.

Поскольку в ТРС хиты симметричны относительно угла ϕ , этот параметр не используется в качестве входных данных. Вместо него рассматривается дополнительный параметр – поперечный импульс p_c . Для обучения NNS используется набор из 500 событий. Размер учебной выборки составляет 75%, тестовой выборки – 15%, валидационной выборки – 10%.

3.5 Разработанные методы

Таким образом, предлагаются следующие методы:

- алгоритм фильтрации треков путем покрытия хитов (HCF);
- алгоритм попарного сопоставления и отбора треков (PWS);
- алгоритм попарного сопоставления и слияния треков (PWM);
- алгоритм отбора треков путем построения графа близости (PGS);
- алгоритм слияния треков путем построения графа близости (PGM);
- алгоритм отбора треков, основанный на нейронной сети (NNS).

3.6 Классификация методов

В таблице 1 представлены рассмотренные алгоритмы разрешения неоднозначностей. Столбцы соответствуют методам кла-



стеризации треков, а строки – методам выбора «лучшего» трека из кластера. Кластеры могут формироваться один раз до начала отбора «лучших треков», а могут, напротив, формироваться динамически во время работы алгоритма. При этом существует два подхода к определению относятся ли треки к одному кластеру:

- подход на основе доли общих точек;
- подход на основе числа общих точек.
- Методы выбора «лучшего» трека из кластера:
- *MaxLength* – «лучшим» треком считается трек, состоящий из большего количества хитов в кластере;
- *NN* – «лучший» трек в кластере выбирается нейросетью;
- *Merge* – «лучший» трек формируется из уникальных точек треков кластера.

Таблица 1. Алгоритмы разрешения неоднозначностей
Table 1. Ambiguity resolution algorithms

	Предварительная кластеризация		Динамическая кластеризация	
	Число общих точек	Доля общих точек	Число общих точек	Доля общих точек
<i>MaxLength</i>	HCF	PGS	—	PWS
<i>Merge</i>	—	PGM	—	PWM
<i>NN</i>	NNS	—	—	—

4. Сравнение методов

4.1 Методика сравнения

4.1.1 Методология исследования

Используется фреймворк MPDRoot [23]. В рамках MPDRoot запускается трекер на основе ACTS [24]. Далее данные из трекера, содержащие информацию о хитах и о треках-кандидатах, выгружаются в файлы. После этого эти файлы подаются на вход алгоритмам разрешения неоднозначностей, которые прототипировались на языке Python⁶.

4.1.2 Тестовые данные

Тестовые данные получаются следующим образом. На первом этапе моделируется генерация частиц с помощью пакета UrQMD, который запускается со следующими параметрами [25]:

- в качестве сталкивающихся частиц принимаются ядра атомов золота $Au + Au$;
- энергия центра масс (*center-of-mass-energy*) принимается равной 7 ГэВ/нуклон;
- параметры удара (*impact parameter*) принимают значения от 0 до 20 фм;
- период, за который генерируются частицы (*calculation time*) принимается равным 200 фм/с, где c – скорость света.

Данные значения соответствуют параметрам эксперимента. Количество событий принимается равным 10 000. На втором этапе методом Монте-Карло симулируется взаимодействие частиц, полученных на предыдущем этапе, и детектора MPD, при этом координата Z вершины взаимодействия принимается равной 0.

⁶ Репозиторий проекта mpd_tpc_tracker [Электронный ресурс] // GitHub, 2024. URL: https://github.com/PlekhanovRUE/mpd_tpc_tracker (дата обращения: 28.08.2024).

4.1.3 Критерий сравнения

Главной характеристикой качества алгоритмов разрешения неоднозначностей является эффективность реконструкции треков, однако эта характеристика не отражает избыточности данных. С другой стороны, при устранении избыточности может теряться информация о треках, что приводит к снижению эффективности.

Наличие ложных треков в множестве распознанных треков способно привести к некорректным выводам о реконструированном событии, поэтому доля ложных треков является второй по значимости характеристикой при оценке качества предлагаемых подходов. Доля ложных треков должна быть близкой к нулю.

При оценке качества алгоритма разрешения неоднозначностей важно принимать во внимание долю треков-дубликатов, поскольку такие треки затрудняют интерпретацию результатов события и требуют значительного количества ресурсов для анализа избыточной информации. Доля треков-дубликатов должна быть минимальной.

4.2 Результаты сравнения

По критериям, описанным в предыдущем разделе, сравниваются шесть алгоритмов, включая уже существовавшие подходы, а также методы, предложенные в данной статье: NNS, HCF, PGM, PGS, PWM, PWS. Имеет смысл рассматривать зависимость показателей качества от псевдобыстроты, поперечного импульса и множественности. На рисунках приведены средние значения характеристик качества алгоритмов. Рассматриваются 10 000 событий.

Наилучшие показатели эффективности после разрешения неоднозначностей имеют треки с псевдобыстротой, находящейся в диапазоне от -1 до 1 (см. рис. 4). При этом лучше всего себя проявляют методы HCF, NNS и PWS. Снижение эффективности при псевдобыстроте в области 0 вызвано наличием ВЧ электрода в детекторе TPC.

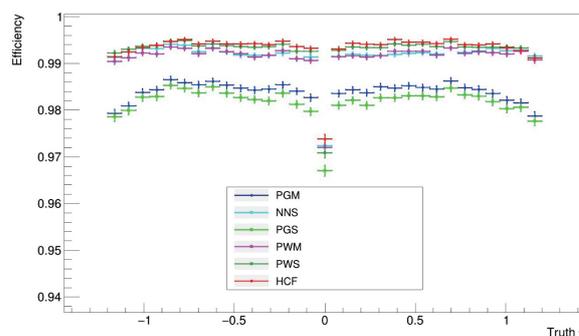
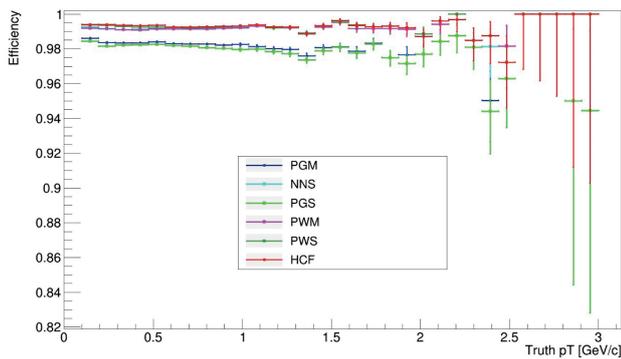


Рис. 4. Эффективность после работы различных методов разрешения неоднозначностей в зависимости от псевдобыстроты

Fig. 4. Efficiency after operation of different ambiguity resolution methods depending on pseudo-fastness

Эффективность реконструкции треков растет с увеличением значения поперечного импульса частицы (см. рис. 5). Показатели эффективности методов разрешения неоднозначностей схожи, при этом лучшие значения показывают HCF, NNS и PWS.

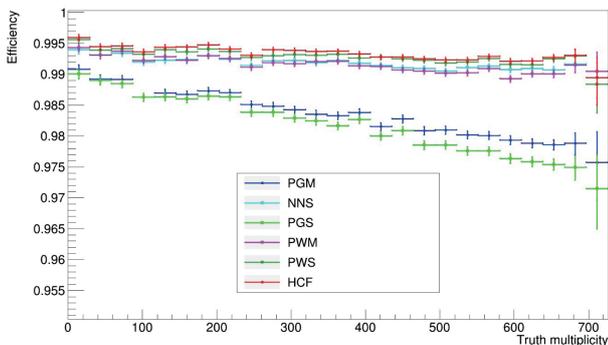




Р и с. 5. Эффективность после работы различных методов разрешения неоднозначностей в зависимости от поперечного импульса

Fig. 5. Efficiency after operation of different ambiguity resolution methods depending on transverse momentum

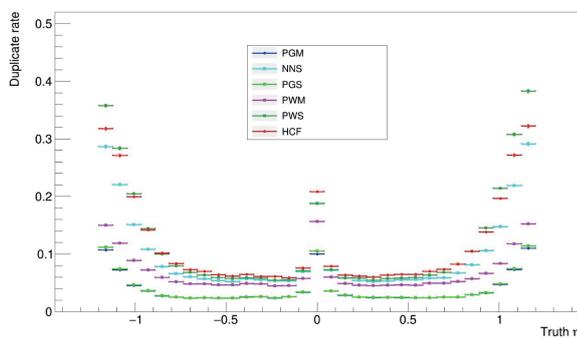
Лучшие результаты в сравнении эффективности в зависимости от множественности показывают методы HCF, NNS и PWS (см. рис. 6).



Р и с. 6. Эффективность после работы различных методов разрешения неоднозначностей в зависимости от множественности

Fig. 6. Post-working performance of different ambiguity resolution methods depending on multiplicity

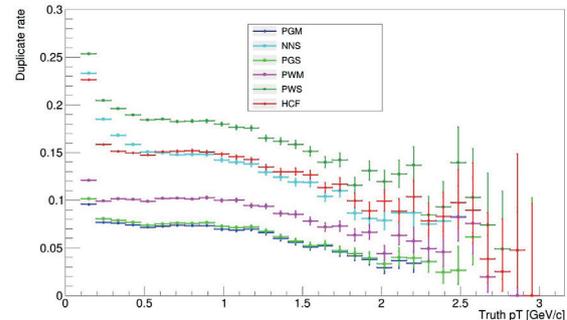
Наиболее низкие значения доли дубликатов имеют треки с псевдобыстротой в диапазоне $0.1 < |\eta| < 0.9$ (см. рис. 7). Среди методов лучше всего себя показывают PGS и PGM, так как они значительно сокращают число треков-дубликатов.



Р и с. 7. Доля дубликатов после работы различных методов разрешения неоднозначностей в зависимости от псевдобыстроты

Fig. 7. The proportion of duplicates after running different ambiguity resolution methods as a function of pseudo-fastness

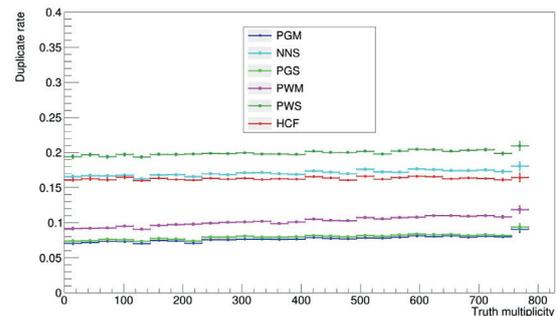
С увеличением поперечного импульса снижается значение доли дубликатов (см. рис. 8). Наиболее низкие значения доли дубликатов показывают методы PGS и PGM.



Р и с. 8. Доля дубликатов после работы различных методов разрешения неоднозначностей в зависимости от поперечного импульса

Fig. 8. The proportion of duplicates after the operation of different ambiguity resolution methods depending on the transverse momentum

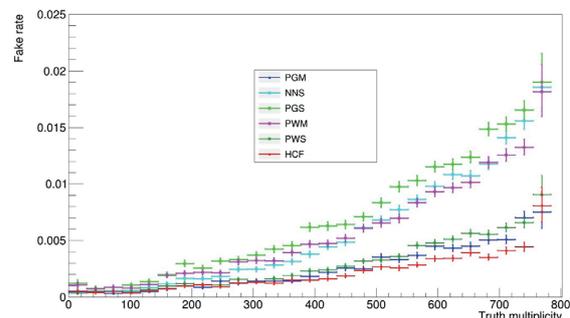
Метод PWS имеет наиболее высокое значение доли дубликатов независимо от множественности (см. рис. 9).



Р и с. 9. Доля дубликатов после работы различных методов разрешения неоднозначностей в зависимости от множественности

Fig. 9. The proportion of duplicates after different ambiguity resolution methods depending on multiplicity

С ростом множественности растет доля ложных треков (см. рис. 10). При этом самые высокие значения доли ложных треков имеет метод PGS. Наименьшие значения доли ложных треков показывают методы HCF, PWS и PGM.



Р и с. 10. Доля ложных треков после работы различных методов разрешения неоднозначностей в зависимости от множественности

Fig. 10. The proportion of false tracks after the operation of various ambiguity resolution methods depending on the multiplicity



Рассмотрим средние значения характеристик качества алгоритмов по всем трекам на 10 000 событий. В столбце RAW представлены характеристики качества трекинга до выполнения алгоритма разрешения неоднозначностей. Из таблицы 2 видно, что метод HCF показал лучшие результаты по двум наиболее приоритетным критериям.

Таблица 2. Значения характеристик качества алгоритмов
Table 2. The values of quality characteristics of algorithms

	RAW	HCF	PWM	PWS	PGS	PGM	NNS
Эффективность, %	99.52	99.28	99.11	99.19	98.34	98.41	99.13
Доля ложных треков, %	00.06	0.11	0.31	0.14	0.36	0.12	0.25
Доля дубликатов, %	97.24	15.57	9.20	18.80	7.29	7.01	16.00

По сформулированным критериям лучшим является алгоритм HCF, так как он наименьшим образом понижает эффективность, оставляет небольшую долю ложных треков. Недостатком алгоритма является сравнительно высокая доля дубликатов.

5. Обсуждение и заключение

В работе были рассмотрены существующие подходы к разрешению неоднозначностей, а также предложены новые алго-

ритмы. Было проведено моделирование 10 000 событий эксперимента MPD методом Монте-Карло. Тестовые данные были использованы для проведения этапа разрешения неоднозначностей предложенными подходами, а также нейросетевым алгоритмом эксперимента ATLAS. Для реализованных подходов были вычислены показатели, предложенные для оценки качества реконструкции треков: эффективность, доля ложных треков и доля дубликатов.

Сравнение подходов показало, что лучшим алгоритмом является HCF с подобранными значениями параметров *newHitsRatio* и *newHitsInRow*. Для эксперимента MPD оптимальными являются значения параметров *newHitsRatio* = 0.25, *newHitsInRow* = 6. Алгоритм с такими параметрами обеспечивает наиболее высокую эффективность реконструкции треков ($\approx 99.28\%$). Этот алгоритм также характеризуется наименьшим временем работы, что важно при работе с большим объемом информации. Похожие результаты могут быть получены и при применении нейросетевого метода эксперимента ATLAS, однако время работы этого алгоритма значительно уступает всем предложенным подходам.

Направлениями дальнейших исследований являются настройка параметров поиска первичных трековых сегментов, параметров поиска треков, а также адаптация наших разработок под новую версию ACTS.

References

- [1] Duckett P, et al. Reconstructing charged particle track segments with a quantum-enhanced support vector machine. *Physical Review D*. 2024;109(5):052002. <https://doi.org/10.1103/PhysRevD.109.052002>
- [2] Burzynski J. Event Reconstruction. In: A Search for Exotic Higgs Decays: Or: How I Learned to Stop Worrying and Love Long-Lived Particles. *Springer Theses*. Cham: Springer; 2023. p. 79-102. http://dx.doi.org/10.1007/978-3-031-30466-8_4
- [3] Abgaryan V, et al. Status and initial physics performance studies of the MPD experiment at NICA. *The European Physical Journal A*. 2022;58(7):140. <https://doi.org/10.1140/epja/s10050-022-00750-6>
- [4] Mankel R. Pattern recognition and event reconstruction in particle physics experiments. *Reports on Progress in Physics*. 2004;67(4):553-662. <https://doi.org/10.48550/arXiv.physics/0402039>
- [5] DeZoort G, et al. Graph neural networks at the Large Hadron Collider. *Nature Reviews Physics*. 2023;5(5):281-303. <https://doi.org/10.1038/s42254-023-00569-0>
- [6] Hough P.V.C. Machine analysis of bubble chamber pictures. In: Proceedings of the International Conference on High Energy Accelerators and Instrumentation. 1959. p. 554-556. Available at: <https://cir.nii.ac.jp/crid/1570854175926719872> (accessed 28.08.2024).
- [7] Ristori L. An artificial retina for fast track finding. *Nuclear Instruments and Methods in Physics Research section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. 2000;453(1-2):425-429. [https://doi.org/10.1016/S0168-9002\(00\)00676-8](https://doi.org/10.1016/S0168-9002(00)00676-8)
- [8] Ososkov G.A., Polyansky A., Puzynin I.V. The modern methods of experimental data processing in high energy physics. *Physics of Elementary Particles and Atomic Nuclei*. 2002;33(3):676-745. Available at: <https://www1.jinr.ru/Pepan/v-33-3/v-33-3-11.pdf> (accessed 28.08.2024). (In Russ.)
- [9] Ju X, et al. Graph neural networks for particle reconstruction in high energy physics detectors. *arXiv:2003.11603*. 2020. <https://doi.org/10.48550/arXiv.2003.11603>
- [10] Dror G., Abramowicz H., Horn D. Vertex identification in high energy physics experiments. *Advances in Neural Information Processing Systems*. 1998;11:868-874. Available at: https://papers.nips.cc/paper_files/paper/1998/file/aff0a6a4521232970b2c1cf539ad0a19-Paper.pdf (accessed 28.08.2024).
- [11] Abidi H, et al. Charged Particle Tracking with Machine Learning on FPGAs. *arXiv:2212.02348*; 2022. doi: <https://doi.org/10.48550/arXiv.2212.02348>
- [12] Batyuk P, et al. Event reconstruction in the BM@N experiment. *EPJ Web of Conferences*. 2019; 204:07012. (In Eng.) doi: <https://doi.org/10.1051/epjconf/201920407012>



- [13] Kalman R.E. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*. 1960; 82(1):35-45. (In Eng.) doi: <https://doi.org/10.1115/1.3662552>
- [14] Frühwirth R. Application of Kalman filtering to track and vertex fitting. *Nuclear Instruments and Methods in Physics Research section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. 1987; 262(2-3):444-450. (In Eng.) doi: [https://doi.org/10.1016/0168-9002\(87\)90887-4](https://doi.org/10.1016/0168-9002(87)90887-4)
- [15] Nikityuk N.M. Methods of processing information from track detectors of high-energy charged particles. *Physics of Elementary Particles and Atomic Nuclei*. 1995; 26(3):719-779. (In Russ.) Available at: http://www1.jinr.ru/Archive/Pepan/1995-v26/v-26-3/pdf_obzory/v26p3_4.pdf (accessed 28.10.2024).
- [16] Baranov D., et al. The particle track reconstruction based on deep neural networks. *EPJ Web of Conferences*. 2019; (214):06018. (In Eng.) doi: <https://doi.org/10.1051/epjconf/201921406018>
- [17] Frühwirth R., Strandlie A. Pattern recognition, tracking and vertex reconstruction in particle detectors. *Springer, Cham*. 2021. 203 p. (In Eng.) doi: <https://doi.org/10.1007/978-3-030-65771-0>
- [18] Zinchenko D., Zinchenko A., Nikonov E. Development of algorithms for track reconstruction and matching in the ITS and TPC detectors at MPD/NICA. *Acta Physica Polonica B Proceedings Supplement*. 2021; 14(3):519-523. (In Eng.) doi: <https://doi.org/10.5506/APhysPolBSupp.14.519>
- [19] Averyanov A., et al. MPD TPC Status. *Physics of Atomic Nuclei*. 2023; 86(5):796-804. (In Eng.) doi: <http://dx.doi.org/10.1134/S106377882305006X>
- [20] Allaire C. et al. Ranking-based neural network for ambiguity resolution in ACTS. *EPJ Web of Conferences*. 2024; 295:03022. (In Eng.) doi: <https://doi.org/10.1051/epjconf/202429503022>
- [21] Aaboud M., et al. Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2. *The European Physical Journal C*. 2017; 77:1-30. (In Eng.) doi: <https://doi.org/10.48550/arXiv.1704.07983>
- [22] Hochreiter S., Schmidhuber J. Long short-term memory. *Neural computation*. 1997; 9(8):1735-1780. (In Eng.) doi: <https://doi.org/10.1162/neco.1997.9.8.1735>
- [23] Busha Ya. Jr., et al. Unified software development and analysis environment for the MPD experiment at the NICA collider. *International Scientific Journal "Modern Information Technologies and IT Education."* 2022; 1:176-182. (In Russ.) doi: <https://doi.org/10.25559/SITITO.18.202201.176-182>
- [24] Ai X., et al. A Common Tracking Software Project. *Computing and Software for Big Science*. 2022. vol. 6. (In Eng.) doi: <https://doi.org/10.1007/s41781-021-00078-8>
- [25] Bleicher M., et al. Relativistic hadron-hadron collisions in the ultra-relativistic quantum molecular dynamics model. *Journal of Physics G: Nuclear and Particle Physics*. 1999; 25(9):1859. (In Eng.) doi: <https://doi.org/10.1088/0954-3899/25/9/308>

Поступила 28.08.2024; одобрена после рецензирования 30.09.2024; принята к публикации 05.10.2024.

Submitted 28.08.2024; approved after reviewing 30.09.2024; accepted for publication 05.10.2024.

Об авторах:

Камкин Александр Сергеевич, ведущий научный сотрудник лаборатории «Гетерогенных компьютерных систем», ФГБОУ ВО «Российский экономический университет имени Г.В. Плеханова (115054, Российская Федерация, г. Москва, Стремянный пер., д. 36); ведущий научный сотрудник Отдела технологий программирования, ФГБУН Институт системного программирования им. В.П. Иванникова РАН (109004, Российская Федерация, г. Москва, ул. А. Солженицына, д. 25), кандидат физико-математических наук, ORCID: <https://orcid.org/0000-0001-6374-8575>, kamkin@ispras.ru

Белецкий Павел Николаевич, научный сотрудник лаборатории «Гетерогенных компьютерных систем», ФГБОУ ВО «Российский экономический университет имени Г.В. Плеханова (115054, Российская Федерация, г. Москва, Стремянный пер., д. 36); программист Отдела технологий программирования, ФГБУН Институт системного программирования им. В.П. Иванникова РАН (109004, Российская Федерация, г. Москва, ул. А. Солженицына, д. 25), ORCID: <https://orcid.org/0000-0003-2072-958X>, belecky@ispras.ru

Бурдельная Вероника Валентиновна, лаборант-исследователь лаборатории «Гетерогенных компьютерных систем», ФГБОУ ВО «Российский экономический университет имени Г.В. Плеханова (115054, Российская Федерация, г. Москва, Стремянный пер., д. 36); аналитик Отдела технологий программирования, ФГБУН Институт системного программирования им. В.П. Иванникова РАН (109004, Российская Федерация, г. Москва, ул. А. Солженицына, д. 25), ORCID: <https://orcid.org/0009-0006-6730-7922>, vv.burdelnaya@ispras.ru

Козмин Илья Александрович, студент кафедры математических методов в экономике Высшей школы кибертехнологий математики и статистики, ФГБОУ ВО «Российский экономический университет имени Г.В. Плеханова (115054, Российская Федерация, г. Москва, Стремянный пер., д. 36); лаборант Отдела технологий программирования, ФГБУН Институт системного программирования им. В.П. Иванникова РАН (109004, Российская Федерация, г. Москва, ул. А. Солженицына, д. 25), ORCID: <https://orcid.org/0009-0001-4752-5280>, lunivilen@ispras.ru

Михалевич Дмитрий Александрович, магистрант научной лаборатории «Прикладное моделирование» Специального факультета талантливой молодежи «Форсайт», ФГБОУ ВО «Российский экономический университет имени Г.В. Плеханова (115054, Российская Федерация, г. Москва, Стремянный пер., д. 36); старший лаборант Отдела технологий программирования, ФГБУН Инсти-



тут системного программирования им. В.П. Иванникова РАН (109004, Российская Федерация, г. Москва, ул. А. Солженицына, д. 25), **ORCID: <https://orcid.org/0009-0009-4758-4510>**, mikhalevich@ispras.ru

Эрконова Джамиля Магомедовна, студент кафедры математических методов в экономике Высшей школы кибертехнологий математики и статистики, ФГБОУ ВО «Российский экономический университет имени Г.В. Плеханова (115054, Российская Федерация, г. Москва, Стремянный пер., д. 36); лаборант Отдела технологий программирования, ФГБУН Институт системного программирования им. В.П. Иванникова РАН (109004, Российская Федерация, г. Москва, ул. А. Солженицына, д. 25), **ORCID: <https://orcid.org/0009-0003-9671-850X>**, erkenovaj@gmail.com

Все авторы прочитали и одобрили окончательный вариант рукописи.

About the authors:

Alexander S. Kamkin, Leading Researcher at the Laboratory of “Heterogeneous Computer Systems”, Plekhanov Russian University of Economics (36 Stremyanny lane, Moscow 115054, Russian Federation); Leading Researcher at the Software Engineering Department, Ivannikov Institute for System Programming of the Russian Academy of Sciences (25 Alexander Solzhenitsyn St., Moscow 109004, Russian Federation), Cand. Sci. (Phys.-Math.), **ORCID: <https://orcid.org/0000-0001-6374-8575>**, kamkin@ispras.ru

Pavel N. Beletsky, Researcher at the Laboratory of “Heterogeneous Computer Systems”, Plekhanov Russian University of Economics (36 Stremyanny lane, Moscow 115054, Russian Federation); Software Developer at the Software Engineering Department, Ivannikov Institute for System Programming of the Russian Academy of Sciences (25 Alexander Solzhenitsyn St., Moscow 109004, Russian Federation), **ORCID: <https://orcid.org/0000-0003-2072-958X>**, belecky@ispras.ru

Veronika V. Burdelnaya, Research Assistant at the Laboratory of “Heterogeneous Computer Systems”, Plekhanov Russian University of Economics (36 Stremyanny lane, Moscow 115054, Russian Federation); Analyst at the Software Engineering Department, Ivannikov Institute for System Programming of the Russian Academy of Sciences (25 Alexander Solzhenitsyn St., Moscow 109004, Russian Federation), **ORCID: <https://orcid.org/0009-0006-6730-7922>**, vv.burdelnaya@ispras.ru

Ilya A. Kozmin, Student at the Chair of Mathematical Methods in Economics, Higher School of Cybertechnology, Mathematics, and Statistics, Plekhanov Russian University of Economics (36 Stremyanny lane, Moscow 115054, Russian Federation); Laboratory Assistant at the Software Engineering Department, Ivannikov Institute for System Programming of the Russian Academy of Sciences (25 Alexander Solzhenitsyn St., Moscow 109004, Russian Federation), **ORCID: <https://orcid.org/0009-0001-4752-5280>**, lunivilen@ispras.ru

Dmitriy A. Mikhalevich, Master Degree Student at the Scientific Laboratory “Applied Modeling”, Department of the Special Faculty of Talented Youth “Foresight”, Plekhanov Russian University of Economics (36 Stremyanny lane, Moscow 115054, Russian Federation); Senior Laboratory Assistant at the Software Engineering Department, Ivannikov Institute for System Programming of the Russian Academy of Sciences (25 Alexander Solzhenitsyn St., Moscow 109004, Russian Federation), **ORCID: <https://orcid.org/0009-0009-4758-4510>**, mikhalevich@ispras.ru

Jamilya M. Erkenova, Student at the Chair of Mathematical Methods in Economics, Higher School of Cybertechnology, Mathematics, and Statistics, Plekhanov Russian University of Economics (36 Stremyanny lane, Moscow 115054, Russian Federation); Laboratory Assistant at the Software Engineering Department, Ivannikov Institute for System Programming of the Russian Academy of Sciences (25 Alexander Solzhenitsyn St., Moscow 109004, Russian Federation), **ORCID: <https://orcid.org/0009-0003-9671-850X>**, erkenovaj@gmail.com

All authors have read and approved the final manuscript.

