

# Обработка экспериментальной информации в новую эру «Больших Данных»

**Ососков Геннадий Алексеевич,**

email: [ososkov@jinr.ru](mailto:ososkov@jinr.ru)

<http://www.jinr.ru/~ososkov>

# Предпосылки

Компьютеры,  
смартфоны  
GPS-навигаторы  
WWW: Google,  
Яндекс,  
вконтакте,  
одноклассники,  
Facebook,  
LinkedIn, ResearchGate  
TV,  
электронные книги

Информация  
Данные  
Картинки  
Аудио  
видео



Бурный рост потоков данных за последние годы, особенно в социальных и бизнес приложениях, где изучаемые явления слишком сложны для их математической формализации, вызвало насущную потребность в выявлении зависимостей **напрямую из огромных, полных информации массивов данных.**

К тому же, только очень малая часть этих данных будет когда-либо востребована, потому что их **объемы слишком велики**, чтобы их вместить в имеющиеся базы данных, а **структура данных слишком сложна или, вообще, не определена** для эффективного их анализа имеющимися алгоритмами за разумное время.

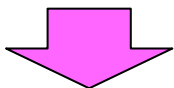
# Много данных

## Пример. ResearchGate

Социальная сеть для научных обменов

Появляется запрос на старую научную работу →

**Сканирование** = оцифрованное изображение



**Распознавание** (FineReader) → **ТЕКСТ**

**Этот процесс – датификация**

Книга превращается в данные!

## Роль технологий

историк Элизабет Эйзенштейн: за 50 лет после Гутенберга

с 1453 по 1503 год было напечатано около восьми

миллионов книг. Это больше, чем все книжники Европы произвели

с момента основания Константинополя примерно 1650 годами ранее

**Сегодня это происходит каждые три дня!**

**2007 год ~ 300 экзабайт** сохраненных данных (7 % - в аналоговом формате, - бумажные документы, книги, фотоснимки и т. д., - остальное в цифровом).

**2013 год - 1,2 зеттабайт** информации, хранящейся в мире (менее 2 % - нецифровая информацию)

**Д. Акст** (книга о будущем электронных библиотек): **годовое поступление** всех мировых источников цифровой информации: оцифрованных книг, фильмов, оптических и магнитных записей требует **полтора петабайта хранения**

ResearchGate

Weekly stats update  
GENNADY ALEXEEVICH OSOSKOV

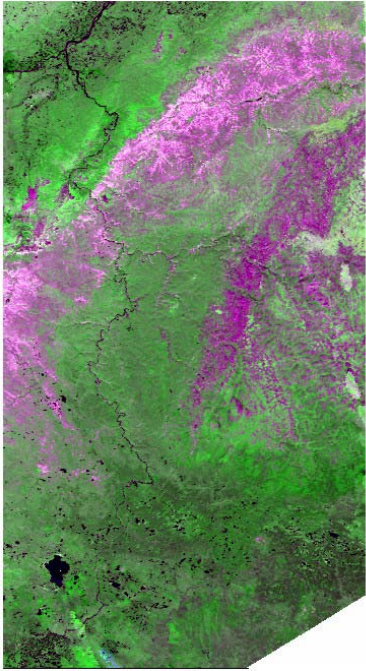
📄 **A new robust fitting algorithm for vertex reconstruction in the CERES experiment**  
Nuclear Instruments and Methods in Physics Research Section A Accelerators Spectrometers Detectors and Associated Equipment 10/1997;   
**52 Downloads**

| Personal stats        | Total | Last week |  |
|-----------------------|-------|-----------|--|
| 🔍 Publication views   | 3k    | 41        | <a href="#">See views by country / institution</a> |
| 📄 Full-text downloads | 772   | 15        | <a href="#">View details</a>                       |
| 👤 Profile views       | 505   | 1         | <a href="#">See views by country / institution</a> |

[See your stats](#)

**1 петабайт=10<sup>15</sup> байт**  
**1 экзабайт=10<sup>18</sup> байт**  
**1 зеттабайт=10<sup>21</sup> байт**

# Примеры больших экспериментальных данных

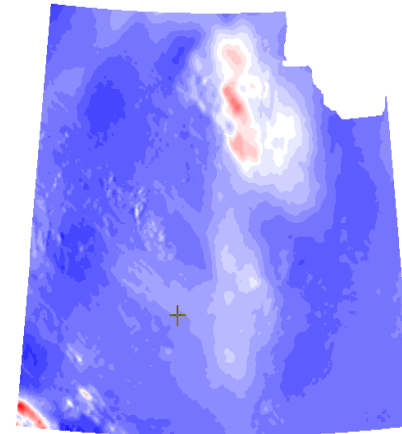
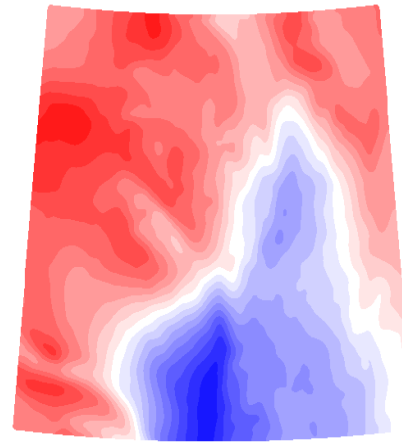


Cosmic pictures of Anabar plato (20 megapixels).

Pictures are shot by 7 channel: 6 in different colors and one in infra-red.

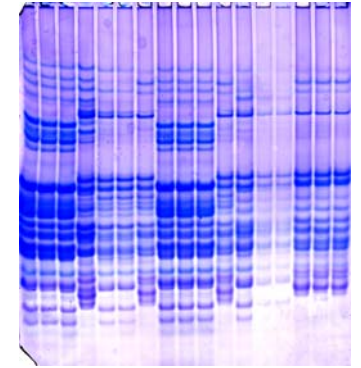
On figures above: left 2 and 3 channels, right channels 1-7.

**Задача:** путем сегментации выделить участки с растительностью, скальные, вспаханные поля.



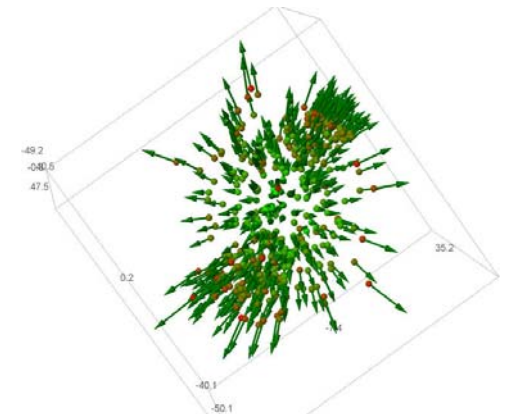
Magnetic (above) and gravitation measurements.

**Задача:** найти геологические структуры с высокой плотностью залегания.



Electrophoregrams of 17 protein specimens (wheat cultivars).

**Задача:** классифицировать по сортам.



Fragments of nucleon-nucleon collision in multifragmentation process. **Задача:** динамическая кластеризация для определения возникающих

# Определения Больших Данных

**Простое определение:** Большие Данные те, что слишком велики и сложны, чтобы их можно было эффективно запомнить, передать и проанализировать **стандартными средствами доступных баз данных и иных имеющихся систем хранения, передачи и обработки.**

Кроме объема, следует учитывать и другие их характеристики. Еще в 2001 году Мета Групп ввела в качестве определяющих характеристик для больших данных так называемые **«три V»:**

- **объём** (*volume*, в смысле величины физического объёма),
- **скорость** (*velocity* в смыслах как скорости прироста, так и необходимости высокоскоростной обработки и получения результатов),
- **многообразие** (*variety*, в смысле возможности одновременной обработки различных типов структурированных и неструктурированных данных)

Однако, когда **общий поток данных растет экспоненциально, удваиваясь каждый год**, за счет революционных технологических изменений, к 2014 году даже эту "3V" модель предлагают расширить, добавляя новые и новые «V»,

- **Validity** (обоснованность, применимость),
- **Veracity** (достоверность),
- **Value** (ценность, полезность),
- **Visibility** (обозримость, способность к визуализации) и т. д.



# Что дают нам Big Data?

**Яндекс** - Андрей Себрант, **Для разогрева**  
Case: магазин Target и беременная девочка (США, 2012)



**Потеря нашей конфиденциальности**

информация о любом нашем обращении в полицию, финансовые налоговые, медицинские учреждения, использование

банковских карт и бонусов сетевых магазинов оседает в соответствующих базах данных и может быть оттуда извлечена по заказу компетентных органов, а также

различных поисковых сетевых служб или, наконец, недобросовестными хакерами.

**Главное другое – большие данные необходимы для выживания в современном мире, т.к. могут дать быстрый прогноз, что делать дальше**



Вредно?

Полезно

# Насколько полезны Big Data ?



Большие должны быть доступны для поисковых систем, проанализированы в центрах бизнеса, производства, медицины, правоохранения, обороны, науки и просто индивидуумов, которые их могут затребовать.

**Что это может дать? – Упростить документооборот, но главное - прогноз!**

**Пример с вирусом H1N1:** Google в 2009 г. точно предсказал распространение гриппа в разных штатах США, проанализировав 3 млрд запросов на лекарства от гриппа на их корреляцию с распространением гриппа во времени и пространстве.

Большие данные помогут решению насущных глобальных проблем, таких как борьба с изменением климата, искоренение болезней, а также содействие эффективному управлению и экономическому развитию. Однако для этого придется пересмотреть традиционные представления об управлении, принятии решений, человеческих ресурсах и образовании.

**Как это делается?**

<http://www.livelib.ru/book/1000755419>

**← читай эту книгу!**

**- Используется не отобранная малая часть данных (выборка), а ВСЁ доступное множество.**

**- Ради этого приходится отказываться от деталей (пренебрегать точностью) и исследовать не причинную связь (почему происходит явление), а корреляционные связи, объясняющие, что именно происходит (полезно, опасно).**

# Data Mining для работы с Big Data

**Data Mining (DM)** - “Это технология, которая предназначена для поиска в больших объемах данных неочевидных, объективных и полезных на практике закономерностей.” *Григорий Пятецкий-Шапиро, 1989 г.*

*Переводы DM на русский:* **добыча** данных, **вскрытие** данных, информационная **проходка**  
**Интеллектуальный анализ данных** — это процесс обнаружения в сырых данных ранее неизвестных, нетривиальных, практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности путем комбинации методов статистики и искусственного интеллекта с использованием технологии баз данных.

В современных условиях данных слишком много, они неоднородны, неполны, неструктурированы и содержат ошибки, а какой-либо рациональной теории для их описания, как правило, нет. Поэтому происходит **сдвиг парадигмы** их обработки **от классической схемы моделирования на основе известной теории**, а затем проверки модели сравнением с экспериментом традиционными средствами анализа данных **к новой парадигме**, когда модели, описывающие связи и зависимости **создаются непосредственно из самих данных н о в ы м и с р е д с т в а м и Data Mining.**

Одно из основных положений Data Mining – поиск **неочевидных** закономерностей. Инструменты Data Mining могут находить такие закономерности самостоятельно и также самостоятельно строить гипотезы о взаимосвязях.



# Задачи, решаемые методами Data Mining

- **Корреляция** - установление статистической зависимости непрерывных выходных от входных переменных.
- **Кластеризация** – это группировка объектов (наблюдений, событий) на основе данных (свойств), описывающих сущность этих объектов. Объекты внутри кластера должны быть "похожими" друг на друга и отличаться от объектов, вошедших в другие кластеры. Чем больше похожи объекты внутри кластера и чем больше отличий между кластерами, тем точнее кластеризация.
- **Классификация** – это отнесение объектов (наблюдений, событий) к одному из заранее известных классов.
- **Ассоциация** – выявление закономерностей между связанными событиями. Примером такой закономерности служит правило, указывающее, что из события X следует событие Y. Такие правила называются ассоциативными.
- **Визуализация**
- **etc**

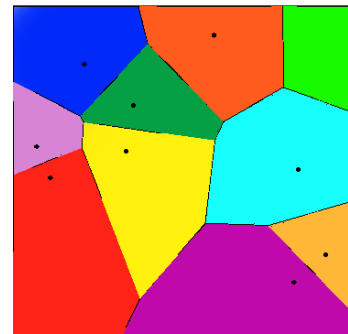
# Кластеризация – важный инструмент Data Mining

Во многих областях современных наук, – биологии, физики, геологии, медицины и т.д. – ученые имеют дело с БОЛЬШИМИ ДАННЫМИ, когда их объем особенно велик ( $10^6$  и больше) . Кроме того, пространство их свойств является многомерным и нет никакой предварительной информации о числе и расположении этих измерений. Поэтому трудно или даже невозможно применить такие известные методы кластеризации, как K-средних.

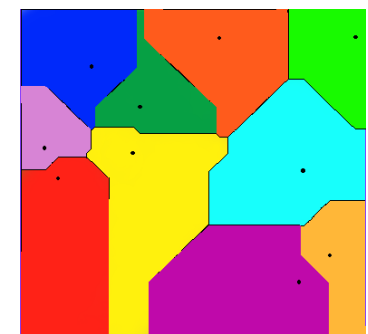
## Новый двухэтапный подход в кластеризации

На первом этапе данные подлежат предварительному объединению, создающему объекты, число которых много меньше числа исходных данных. Для такой предварительной кластеризации (или квантизации) мы выбрали **разбиение Вороного**. Оно разбивает исходное векторное пространство на множество выпуклых многогранников так, что любая точка такого многогранника лежит ближе к его центру, чем к центру любого из остальных многогранников.

**Пример.** Оценка районов принадлежности покупателей к ближайшему магазину (метод ближайшего соседа).  
Когда покупатель идет до магазина пешком, срезая углы, он движется по Эвклидовой метрике, но если он едет на машине в соответствии с направлениями улиц, то правильное описание его пути – манхеттенское расстояние



Voronoi cells for  
Euclidean distance



Voronoi cells for  
Manhattan distance

# Связь триангуляции Делоне и Разбиений Вороного

Триангуляция Делоне однозначно соответствует диаграмме Вороного: триангуляция связывает центры тех областей Вороного, которые имеют общие границы.

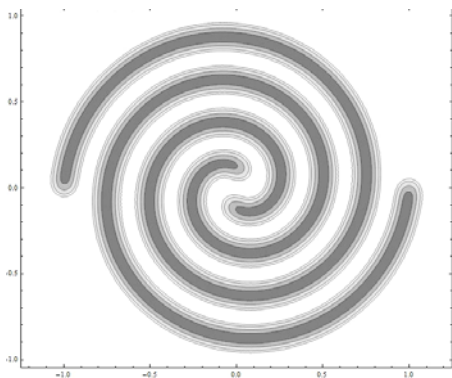
**Однако, к примеру, для квантизации изображений задача является обратной, триангуляция Делоне неизвестна, а требуется найти центры мозаики Вороного**

## Как это работает методом Растущего Нейронного Газа

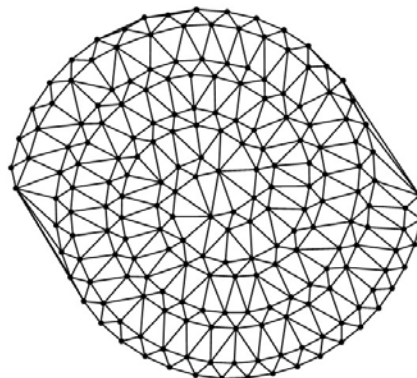
*S.V. Mitsyn, G.A. Ososkov, The Growing Neural Gas and Clustering of Large Amounts of Data, Optical Memory and Neural Networks (Information Optics), 2011, Vol. 20, No. 4, pp. 260–270*

An example of objects to be partitioned into Voronoi mosaic

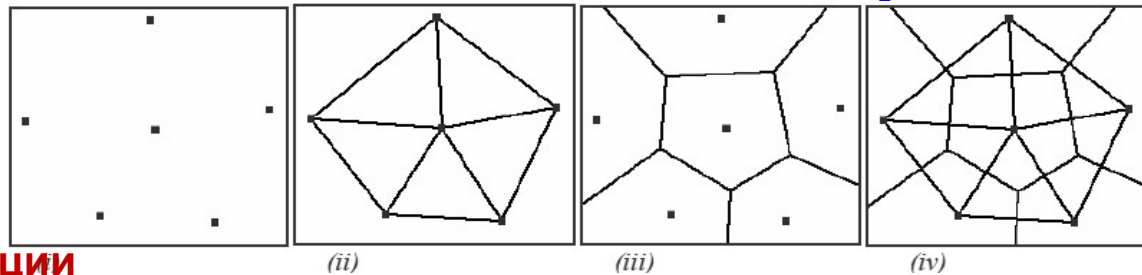
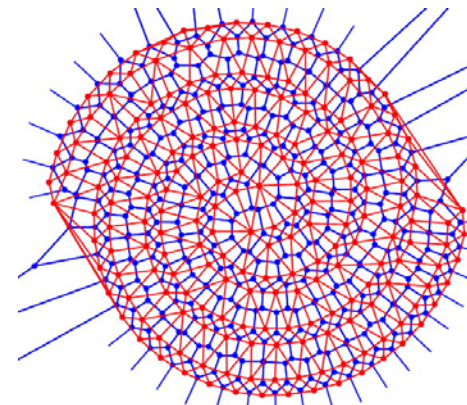
source data (2 million points)



Delaunay triangulation



Voronoi mosaic (100 cells only)

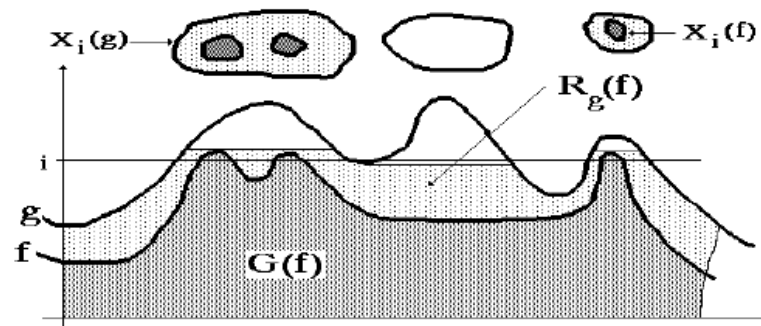
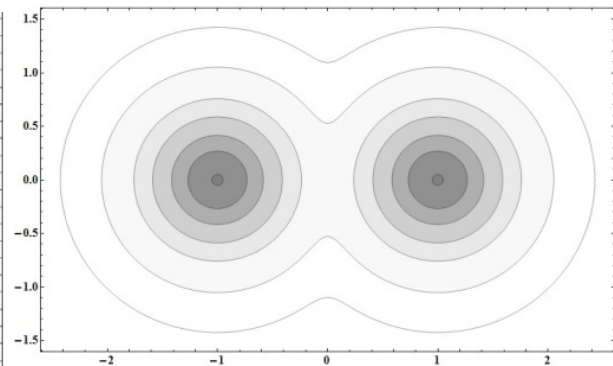
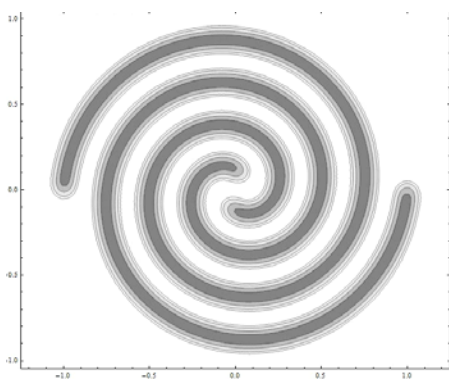


Formation of a Voronoi diagram on a plane: (i) nodes on the plane, (ii) Delaunay triangulation, (iii) Voronoi diagram, (iv) superposition of the Delaunay triangulation and the resulting Voronoi diagram.

# Второй этап кластеризации

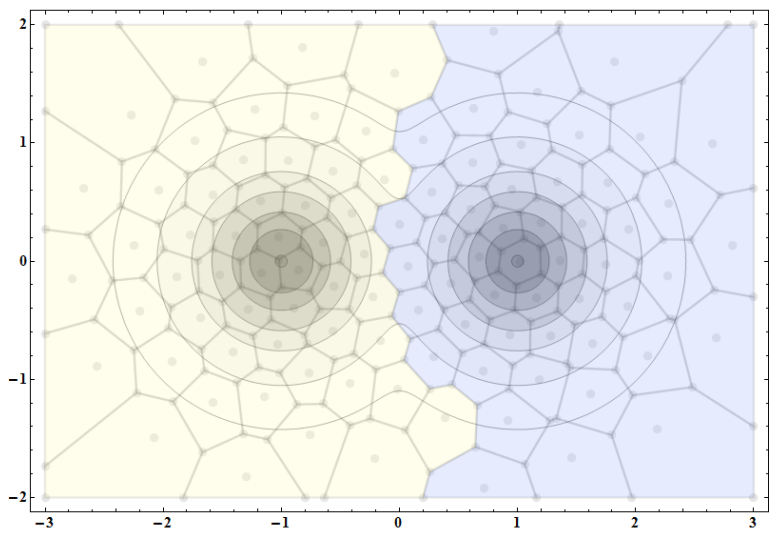
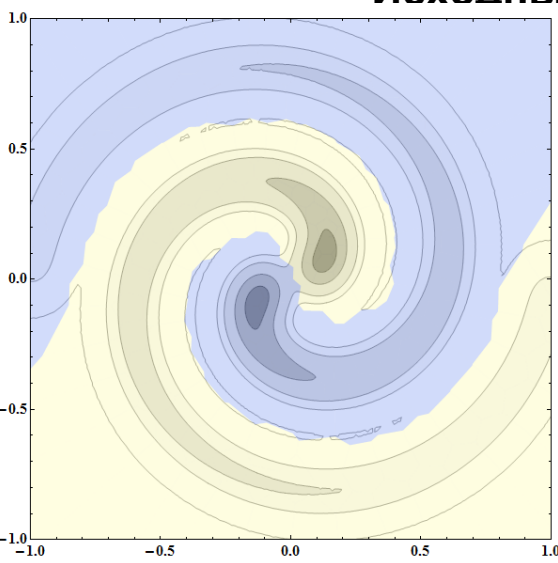
много известных алгоритмов: ближайший сосед, метод Варда и т.д.

## Новый метод- алгоритм водораздела



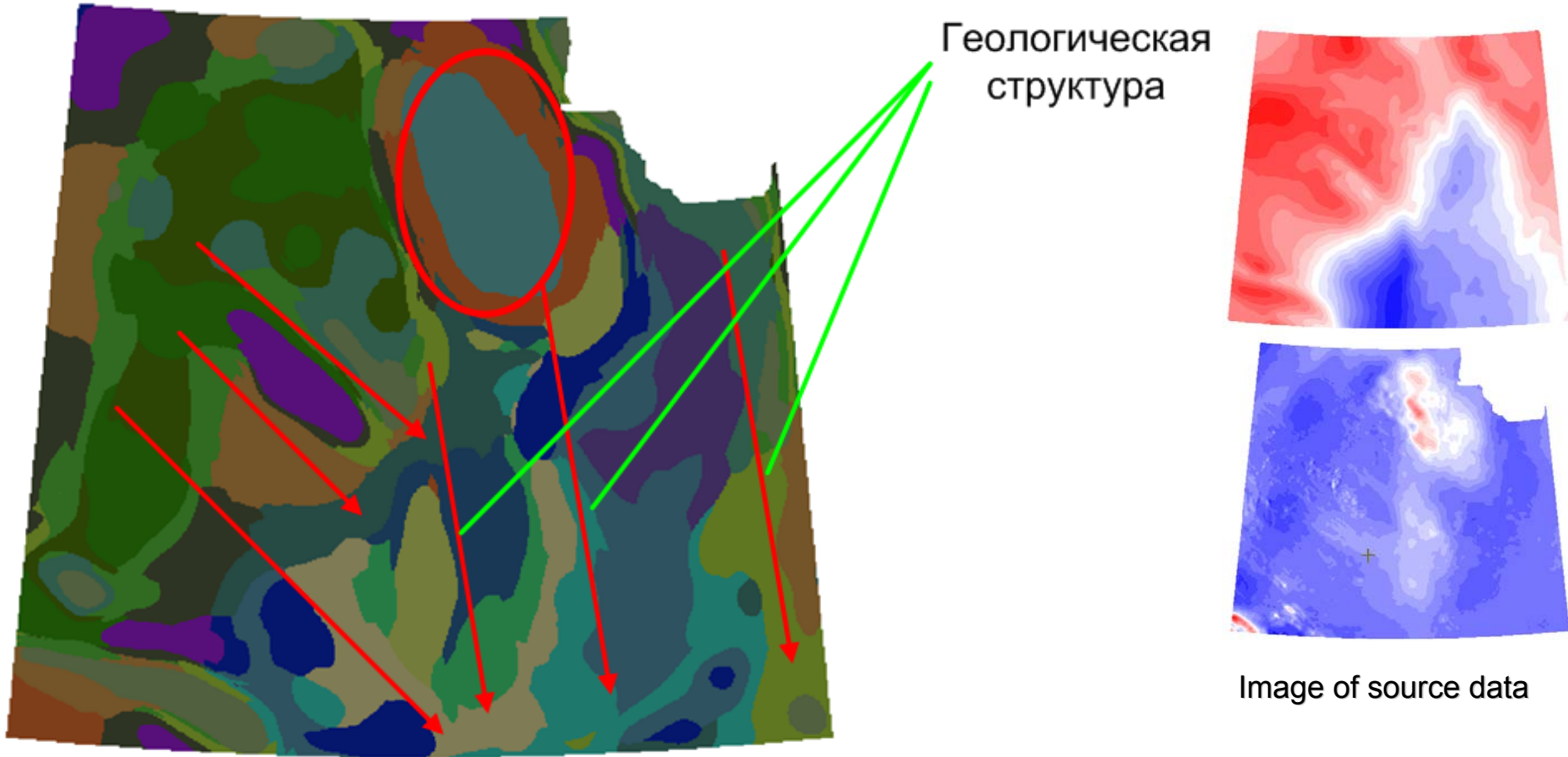
Водораздел как геодезическая реконструкция

Исходные множества точек



Results of watershed clustering. Thanks to Sergey Mitsyn

# Пример: кластеризация геофизических данных





# Методы Data Mining

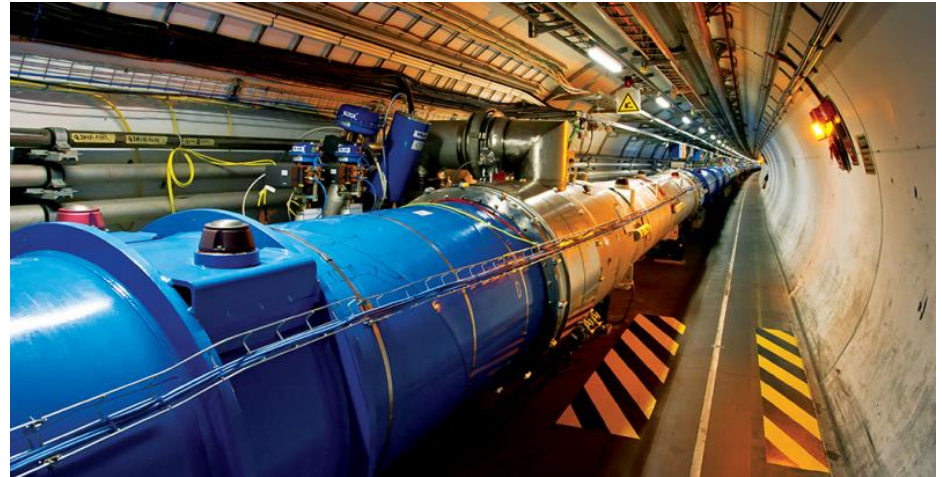
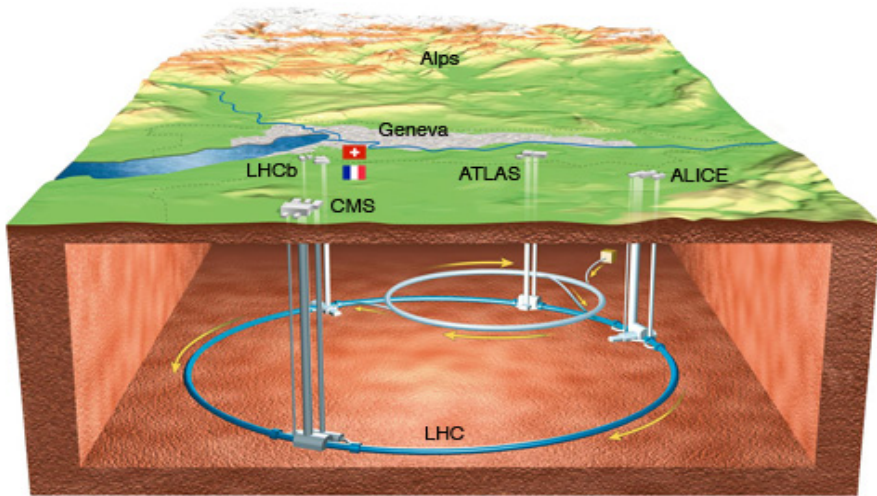
- **Непосредственное использование данных:** дескриптивный анализ и описание исходных данных, кластерный анализ
- **Статистические методы:** анализ связей (корреляционный и регрессионный анализ, факторный анализ, дисперсионный анализ). Анализ временных рядов (динамические модели и прогнозирование).
- **Методы искусственного интеллекта:** искусственные нейронные сети (распознавание, кластеризация, принятие решений, прогноз); генетические алгоритмы; ассоциативная память (поиск аналогов, прототипов); нечеткая логика; деревья решений; системы обработки экспертных знаний.

Хотя методы DM ориентированы главным образом на «майнинг» в бизнесе и социальных науках, они также часто применяются и в таких технических и научных областях, как биоинформатика, генетика, медицина, образование и электроэнергетические расчеты. Существует большой объем прикладных DM-программ в открытом, но больше – в коммерческом доступе.

Однако **DM-приложений для экспериментальной физики вы там не найдете.**

*Чтобы понять, почему - посмотрим на некоторые эксперименты в физике высоких энергий (ФВЭ) и то, какие данные в них получают*

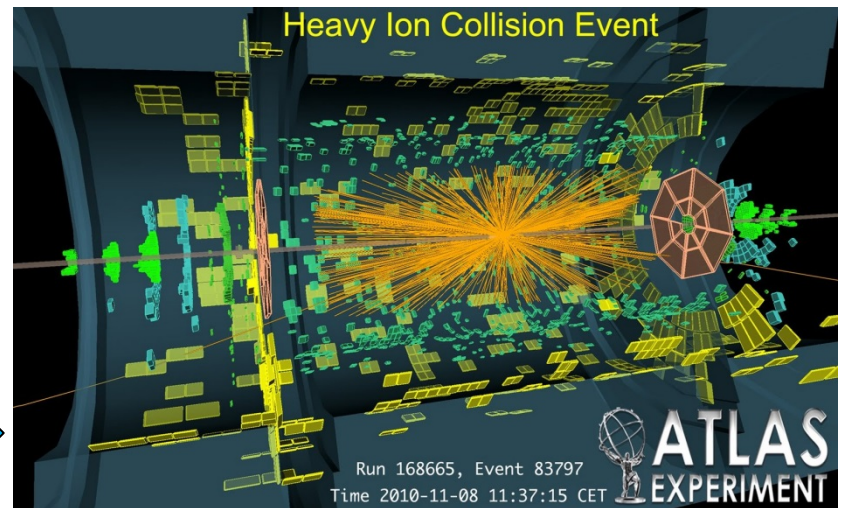
# Эксперименты ФВЭ и ядерной физики



## Знаменитый БАК - Большой Адронный Коллайдер в ЦЕРНе



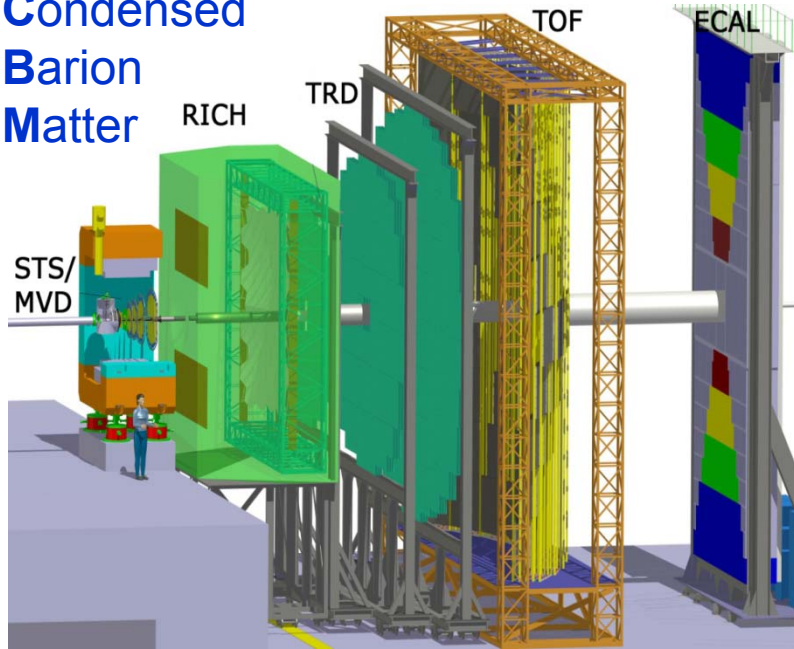
2012  
Получен  
бозон  
Хиггса





# Эксперименты в Германии и Италии

Condensed  
Barion  
Matter



schematic view of the CBM setup



1600 m in depth  
~100'000 m<sup>3</sup> caverns' volume



Эксперимент **OPERA**: поиск осцилляций нейтрино

**CBM experiment** (Germany, GSI,

to be running in 2018)

10<sup>7</sup> events per sec,

~1000 tracks per event

~100 numbers per track

**Total: terabytes/sec !**

# Анализ данных в детекторах СВМ

## Реконструкция событий

$10^7$  соб/сек,  $\sim 1000$  треков/соб,  $\sim 100$  чисел/трек

в режиме on-line!

### 1. Вершинный детектор STS

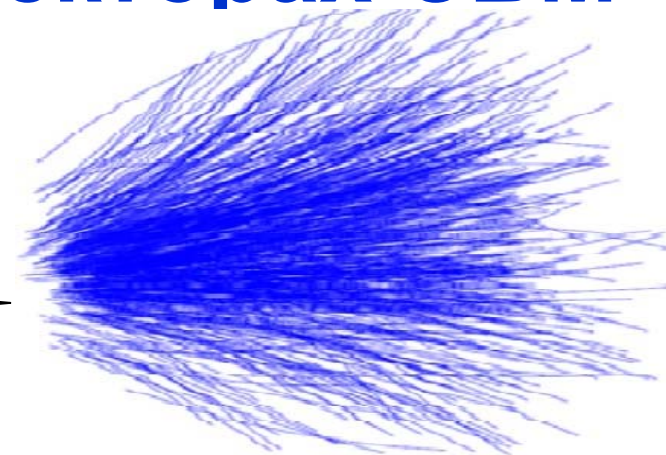
Задачи трекинга в магнитном поле:

- Распознавание треков
- Вычисление их параметров
- Определение вторичных вершин для короткоживущих и нейтральных частиц

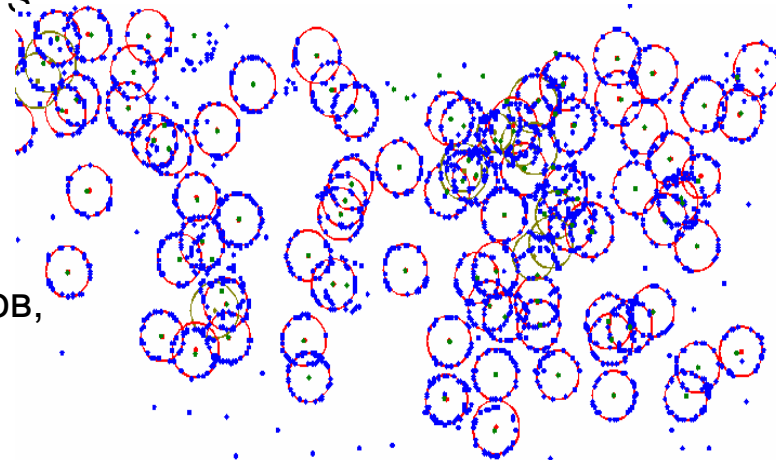
### 2. Детектор RICH

Задачи по идентификации электронов и пионов:

- распознать все кольца
- вычислить их параметры с учетом шумовых отсчетов, перекрытия колец и оптических искажений, из-за которых кольца выглядят, как эллипсы
- стыковать кольца с треками, найденными в STS
- определить, какие треки порождены электронами



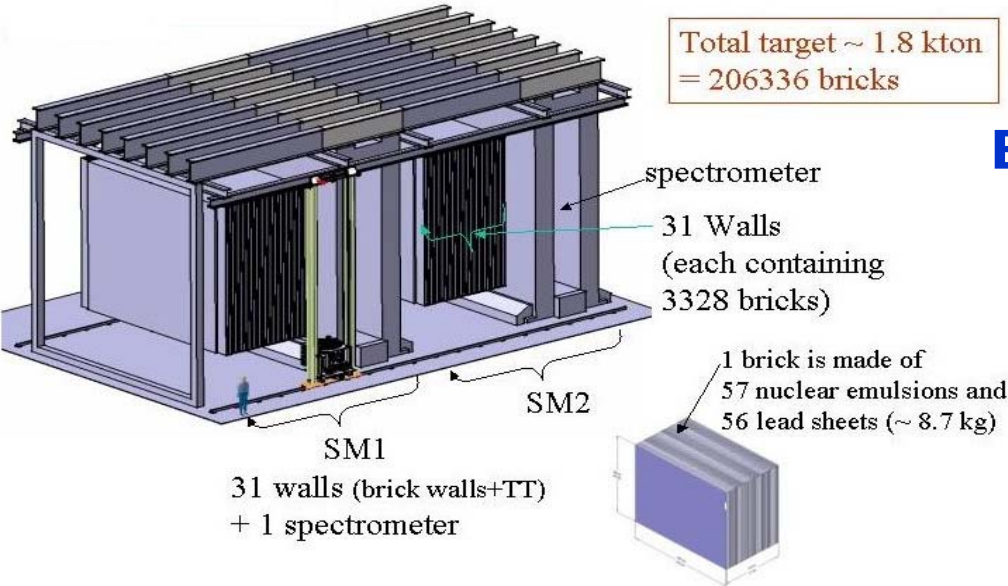
Модельный вид Au+Au взаимодействия в STS



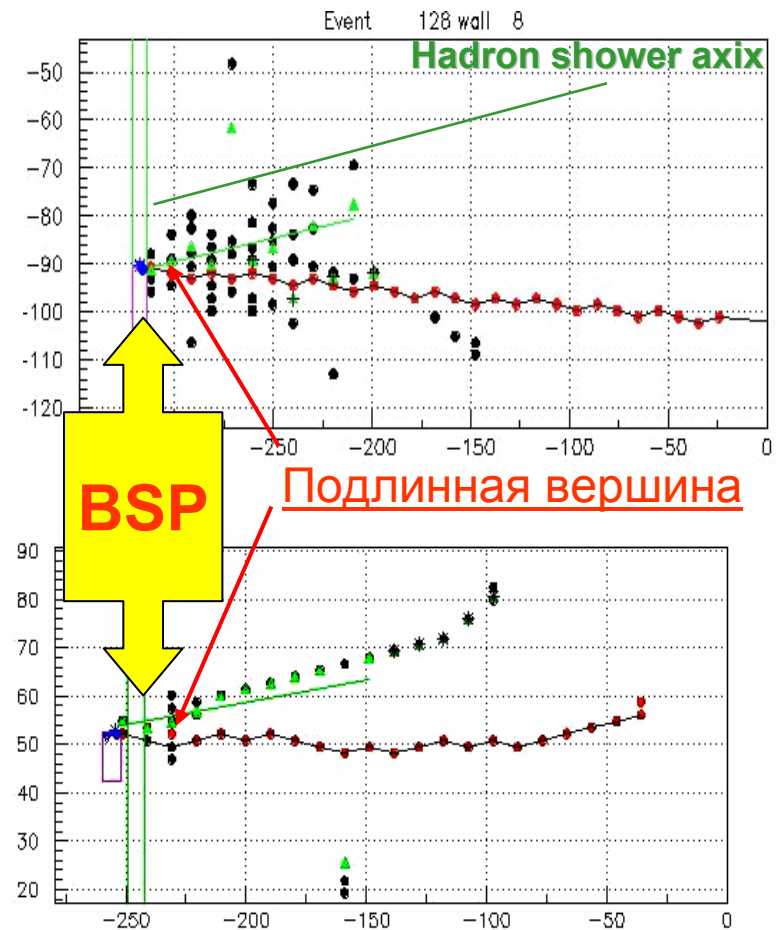
Кольца черенковского излучения, зарегистрированные в детекторе RICH

## В эксперименте OPERA

Мы не рассматриваем вопросы сканирования эмульсии для поиска осцилляций нейтрино, т.к. это – отдельная задача



Трекер образован 31 стеной из эмульсионных кирпичей, перемежаемых решетками из 7-миллиметровых сцинтилляторов, снабженных электронными регистраторами. **Основная задача** анализа данных трекера – **определение того конкретного кирпича, где произошло искомое событие осцилляции нейтрино**. Треки прослеживаемые по точкам пересечения сцинтилляторов, засвеченных проходящими частицами, должны сходиться в вершину, которая и определяет искомый кирпич. Главная трудность - частицы обратного рассеивания (**back-scattered particles – BSP**) случающиеся в 50% событий и не содержащие полезной информации



Два типа событий OPERA с BSP



# Отличия DM и анализа данных в ФВЭ и ЯФ

различие

Физики, захлестываемые потоками данных от экспериментов и моделирования физических процессов, разработали **свой собственный всеохватывающий набор методов анализа данных (Data Analysis – DA)**, реализованный в известной программной платформе **ROOT**, на которой теперь основаны почти все программные оболочки – фреймворки большинства европейских экспериментов. В отличие от DM, методы DA в физике высоких энергий и ядерной физике используют выдающиеся достижения теоретической физики, дающие возможность успешно **моделировать сложнейшие физические процессы, происходящие в экспериментальных установках** при взаимодействиях частиц в каждом из детекторов и траекторий получившихся осколков в каждом из компонентов этих детекторов с учетом их материалов и магнитных полей.

общее

Методы DA – это только часть общего гигантского процесса манипулирования данными в современных экспериментах ФВЭ и ЯФ. Помимо задач анализа данных не менее значительную часть занимают **задачи хранения и обмена данными** в иерархической **ГРИД-облачной системе распределенных вычислений**, объединяющей Tier-центры разных уровней.

**Концепция использования распределенных облачных систем для хранения, распределения и обработки данных является общей для физиков и бизнесменов**

# Этапы процессов DA в ФВЭ и ЯФ 1

Важнейший этап – предобработка включает

- **Получение и сохранение данных:** до применения алгоритмов DA данные, подлежащие исследованию должны быть зарегистрированы, преобразованы из отсчетов детекторов в формат обычных единиц измерений;
- **Селекция данных:** фильтрация от шума и несущественных измерений, не удовлетворяющих заданным условиям. Проверка этих условий выполняется системой «умных» триггеров разных уровней и ведет к сокращению объема данных на много порядков;
- **Преобразование данных (калибровка и алайнмент)** для перевода в формат подходящий для последующего анализа и хранения.

# Этапы процессов DA в ФВЭ и ЯФ 2

Следующие этапы можно суммировать как

■ **Распознавание образов для реконструкции событий**: трекинг, нахождение вершин событий, распознавание колец черенковского излучения, а также выявление и удаление ложно распознанных объектов. Применяемые методы

- преобразования Хафа,
- клеточные автоматы,
- фильтр Калмана,
- искусственные нейронные сети,
- вейвлет-анализ и др.

■ **Оценивание физических параметров**

- методы математической статистики;
- робастное оценивание

■ **Проверка гипотез**

- отношения правдоподобия,
- нейросетей
- усиленные алгоритмы машинного обучения (boosted decision trees - BDT).

*Ниже будут даны примеры применения методов DA из практики ОИЯИ*

□ **Моделирование выполняется на всех этапах анализа данных**

# Этапы процессов DA в ФВЭ и ЯФ 3

Следует подчеркнуть очень важную **роль моделирующих программ** (таких как **GEANT**) на всех этапах анализа данных.

**Моделирование позволяет:**

- Оптимизировать по деньгам, материалам и времени всю экспериментальную установку и разработать алгоритмы DA еще на стадии проектирования;
- Разработать и протестировать необходимую программную оболочку эксперимента;
- Оптимизировать структуру и необходимое оборудование запланированных детекторов, минимизируя стоимостные и временные затраты при заданной эффективности и точности работы детектора;
- Рассчитать заранее все необходимые распределения, пороги для проверки гипотез и сгенерировать обучающие выборки для искусственных нейронных сетей.

# Обработка данных в экспериментах БАК

БАК в 2012 г.

ALICE,  
ATLAS,  
CMS  
LHCb

Сокращение  
в миллион  
раз!

Вспомните Д.Акста: 1,5  
петабайта – годовое хранение!

один петабайт  
=  $10^{15}$  байт  
данных/сек

Запомнить такое количество  
данных невозможно ни на  
какой из современных  
вычислительных систем

Система триггеров разных уровней выполняла  
сверхбыструю сложную электронную предобработку,  
оставлявшую только одно полезное физическое событие  
из 10 тысяч

Дальнейший анализ выполнялся в ЦЕРНовском компьютерном  
центре обработки из многих тысяч процессоров. Оставалось только  
1% событий, возможно содержащих искомый физический феномен.

**Идея распределенной обработки**

**Итог - 25 петабайт данных в год**, которые требовалось хранить в специальных  
роботизированных ленточных хранилищах, т.к. копии этих данных подлежат передаче  
в сотни физических центров 36 стран мира для более тщательного анализа



# Всемирная сеть распределенных вычислений - WLCG

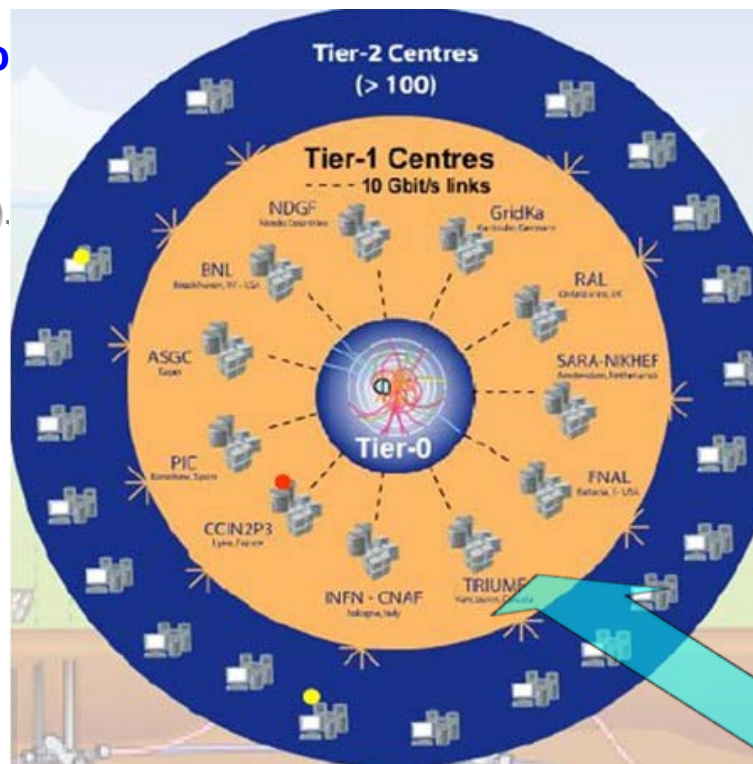
Сотни тысяч компьютеров, объединены во **Всемирную сеть распределенных вычислений** - Worldwide LHC Computing Grid (WLCG).

**Иерархическая структура WLCG состоит из вычислительных Tier-центров разных уровней.**

Ежедневно в WLCG обрабатываются полтора миллиона заданий, на что одному даже самому мощному современному компьютеру потребовалось бы 600 лет.

Помимо задач анализа данных значительную часть занимают **задачи хранения и обмена данными в системе WLCG**

- Создание баз наблюдаемых и смоделированных данных и хранение их копий (реплик)
- Распространение и обмен репликами по запросам WLCG центров разных уровней.
- Сбалансированный процесс копирования востребованных и стирания устаревших записей



|                          |  |
|--------------------------|--|
| Tier0<br>CERN            | Первичная реконструкция событий, калибровка, хранение копий полных баз данных  |
| Tier1<br>11<br>центров   | Полная реконструкция событий, хранение актуальных баз данных по событиям, создание и хранение наборов анализируемых событий, моделирование, анализ |
| Tier2<br>>200<br>центров | Репликация и хранение наборов анализируемых событий, моделирование, анализ   |

**В 2014 г. CMS Tier-1 центр начнет работать в ОИЯИ**

# Наука с интенсивной обработкой данных – новая парадигма

На торжестве 4 июля 2012 г. по поводу получения ЦЕРНом нобелевской премии за открытие бозона Хиггса директор ЦЕРНа Рольф Хойер прямо назвал GRID-технологии одним из трех столпов успеха (наряду с ускорителем LHC и физическими установками).

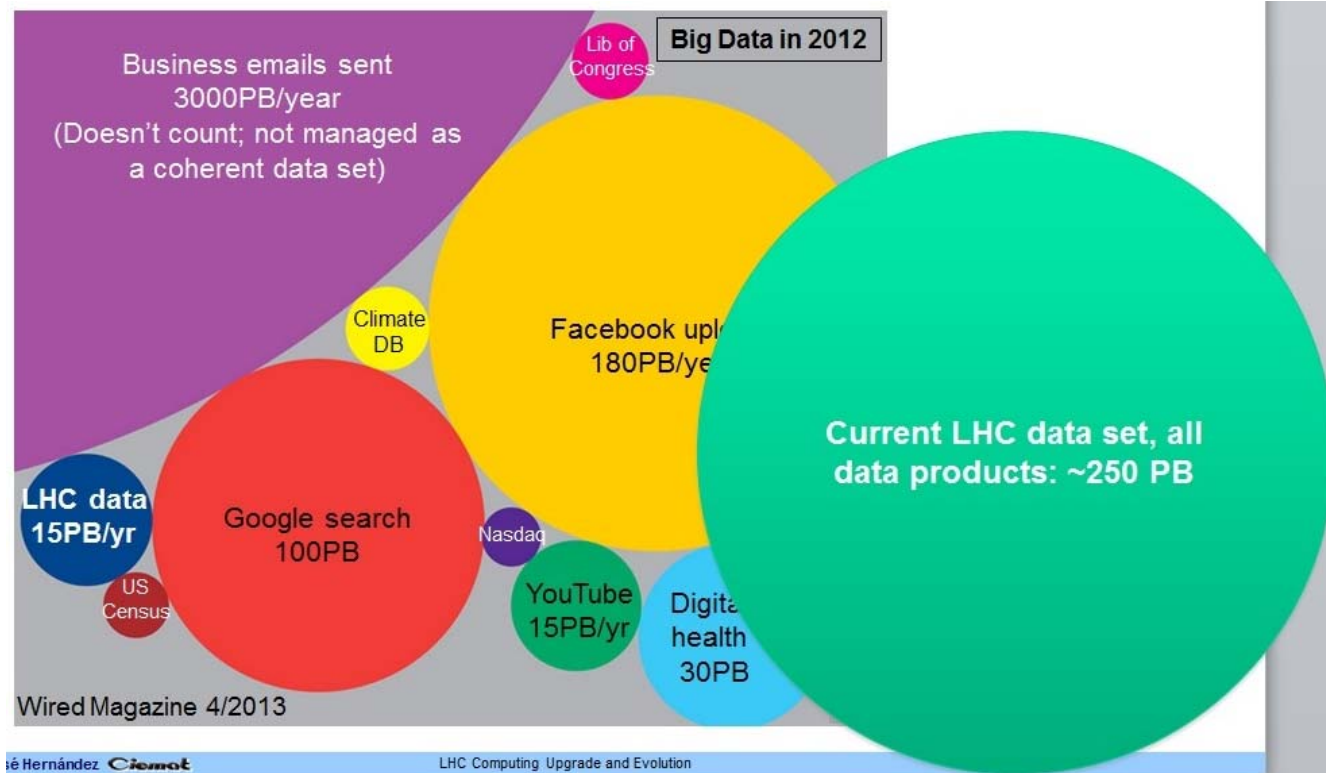


Этот успех также подтверждает, что **ЦЕРН входит в эру Больших Данных** и эффективно преодолевает проблемы четвертой парадигмы, что является одним из примеров (наряду с созданием в ЦЕРНе WWW-всемирной паутины), когда разработки в области физики частиц начинают влиять на исследования в других научных областях.

Известный специалист Microsoft в области хранения информации Джим Грэй предсказал, в 2005 г., что вступление научных исследований в эпоху пета- и экза-данных должно неизбежно потребовать развития новой **науки с интенсивной обработкой (Data-intensive science)** и назвал это изменение **«четвертой парадигмой науки»**, в дополнение к трем предыдущим научным парадигмам — экспериментальной, теоретической и вычислительной

# Вступление ФВЭ и ЯФ в эру Big Data

Сравнительная диаграмма по общим объемам перерабатываемых в 2012 году данных в социальных сетях, поисковых системах, разных отраслях бизнеса, медицины, климатических прогнозов и БАК наглядно показывает, что исследования в ЦЕРНе идут в **условиях Больших Данных**.



Более того, вторичный запуск БАК, предстоящий в 2015 году, потребует его существенной модернизации, когда поток данных возрастет в 2,5 раза при удвоении времени на их обработку.

# Компьютинг ФВЭ в эпоху Больших Данных

Планы развития компьютеринга в ЦЕРНе для обеспечения **потенциально новой физики** после запуска БАК в 2015 году

1. Значительное увеличение вычислительных мощностей и сетевых ресурсов хранения данных;
2. Развитие интеллектуальных средств динамического хранения данных;
3. Повышение эффективности WLCG путем синтеза грид и облачных технологий;
4. активизация использования распределенных параллельных вычислений
5. Совершенствования алгоритмических и программных средств анализа и моделирования;

Такие же планы характерны для ведущих физических центров мира, в том числе и для ОИЯИ с его мегапроектом NICA.

*Рассмотрим их более детально*

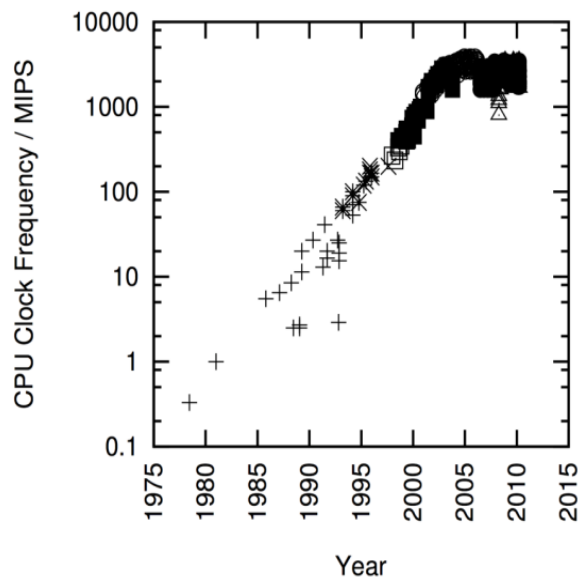


# 1. Значительное увеличение вычислительных мощностей и сетевых ресурсов хранения данных

## Большие данные =

- больше CPU (по стоимости на 15% в год)
- рост дискового пространства (на 15%)
- рост роботизированных библиотек массового хранения (на 15%)

цифры касаются T0 в ЦЕРНе и T1-T2 в странах, входящих в WLCG



Важный нюанс: **Закон Мура**: экспоненциальный рост числа транзисторов по годам = соответствующий рост числа производимых CPU, однако скорость их работы сдерживается из-за **эффекта «тепловой смерти»**.

**Поэтому для увеличения производительности вычислительных процессоров потребуются**

- вычислители с большим числом ядер для введения параллелизма и/или
- применение новых графических GPU процессоров

После 2014 года более 15% данных LHC будут обрабатываться в российских центрах WLCG.

## 2. Интеллектуальные средства динамического хранения данных 1

- Терабайты данных в секунду, производимые в экспериментах ФВЭ и ЯФ, требуют либо сложнейших многоуровневых триггерных процедур или сверхбыстрой параллельной обработки данных для сжатия сырых данных в миллионы раз.
- **Итог - 25 петабайт данных в год,** которые требовалось хранить в специальных **роботизированных ленточных хранилищах**, т.к. копии этих данных подлежат передаче в сотни физических центров 36 стран мира для более тщательного анализа. Для архивирования, изготовления и передачи копий разработаны сложные роботизированные хранилища ленточных картриджей, каждый из которых вмещает по 6.25 терабайт

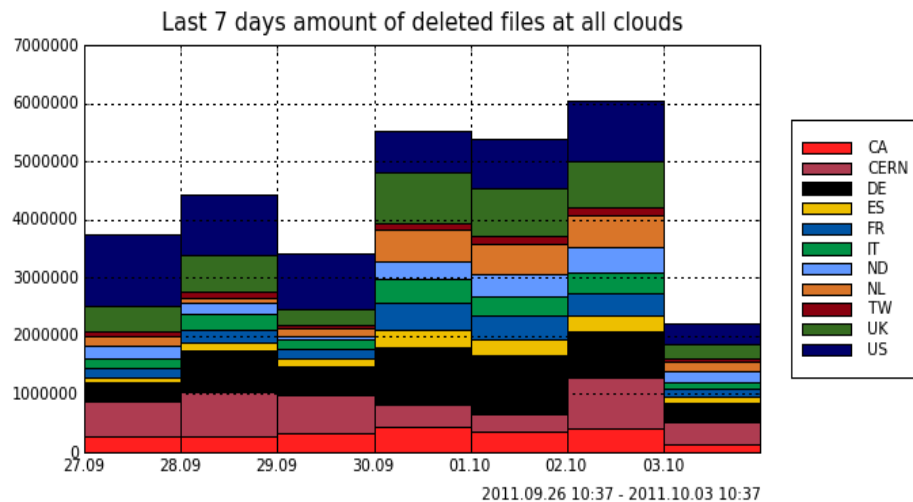




# Интеллектуальные средства динамического хранения данных 2.

- Одной из ключевых проблем в такой системе управления распределением данных является **сбалансированный процесс копирования востребованных (горячих) и стирания устаревших (холодных) записей.**
- Проект ATLAS Distributed Data Management - пример автоматической организации такого процесса, выполняющего копирование, доступ, стирание и весь учет оборота данных эксперимента ATLAS в более, чем 120 грид-сайтах. Разработан специальный сервис учета распространения данных (data popularity service) для принятия решений о необходимости увеличения или уменьшения числа копий или полного стирания файла с данными, как вышедшего из употребления.

В ЛИТ ОИЯИ разработана и установлена программа ATLAS Deletion Service DQ2, которая ежедневно находит и стирает 2-2,5 миллиона файлов, освобождая до 500 терабайт (порядка 300к файлов в час).



# 3. Повышение эффективности WLCG путем синтеза грид и облачных технологий

Жесткая структура системы грид создавалась для интеграции уже существующих аппаратных и программных ресурсов, зафиксированных в системе, в то время как облачная структура распределенных вычислений оказывается более гибкой, используя виртуальные кластеры из виртуальных вычислителей.

Николай Кутовский, диссертация: *Включение в грид облачных структур позволяет сократить время решения широкого круга задач в области физики высоких энергий и повысить эффективность использования ресурсов*

Примером уже имеющейся технологии, реализующей подобный синтез для работы с Большими Данными является система PanDA (Production and Distributed Analysis – обработка данных и распределенный анализ) эксперимента ATLAS на LHC. Сегодня PanDA развилась в систему BigPanda и уже работает для российского мегапроекта NICA в ОИЯИ (доклад А.Климентова на этой школе).

## За большими данными следит ПАНДА

Текст  
А. Климентов, сотрудник Брукхавенской национальной лаборатории,  
А. Ваншин, сотрудник Аргонской национальной лаборатории,  
В. Кореньков, ОИЯИ, г. Дубна  
Иллюстрация Владимир Камнев



Броский и звучный термин «Big Data» прочно занял свое место в лексиконе не только ученых и компьютерных профессионалов, но и политиков разного ранга.

Суперкомпьютеры №15, 2013, стр.56

# Моделирование грид-облачных систем

- Разработка сложнейших грид-облачных систем сбора, передачи и распределённой обработки сверхбольших объемов информации требует больших предварительных исследований по выбору **оптимальной их структуры с учетом стоимости и предполагаемых ресурсов и загрузки.**
- Подобные исследования должны основываться на **тщательном моделировании** как потока заданий с учетом их типов и статистических данных о распределении времени их поступления и требуемых компьютерных ресурсов для их выполнения, так и состава моделируемой грид-структуры.
- Такая программа моделирования SyMSim (Synthesis of Monitoring and Simulation) разработана в ЛИТ ОИЯИ для ускорения проектирования центров WLCG Tier 0 NICA и Tier 1 CMS. Программа ориентирована на повышения эффективности их разработки путем учета качества работы уже функционирующей системы в прогнозах на ее дальнейшее развитие. Это выполнено за счет **объединения самой программы моделирования с системой мониторинга реального (или модельного) грид-облачного сервиса** через специальную базу данных, осуществляющую сбор и статистический анализ по вычислению распределений данных мониторинга, используемых затем для динамической коррекции параметров моделирования.

# Модель облачной структуры хранения данных

Рассматривается модель реализации облачной структуры, предназначенной для хранения данных в роботизированной библиотеке с тысячами кассет с магнитными лентами, которые робот автоматически достаёт и устанавливает в одно или несколько устройств чтения-записи (драйвов).

## Проектируемая структура:

ленточный робот, массив драйвов, кластер процессоров.

**Стоимость** драйва - 5 условных единиц, процессора - 3 единицы.

**Критерий оценки:** время прохождения тестового потока из 100 заданий.

**Бюджет:** 100 условных единиц

**Требования к проекту:** оптимальное соотношение количества процессоров и количества драйвов в пределах бюджета

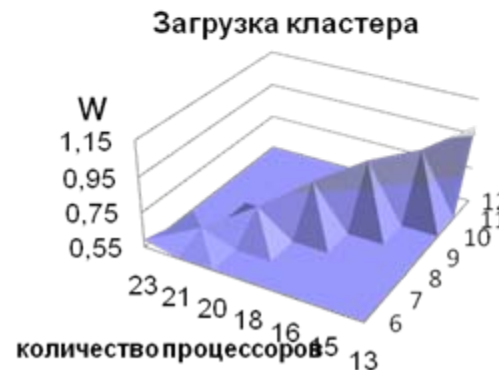
## Результаты моделирования.

1. *Определение степени загрузки кластера*

Загрузка кластера  $W = T_{100}/T_a$ , где

$T_{100}$  – процессорное время выполнения пакета

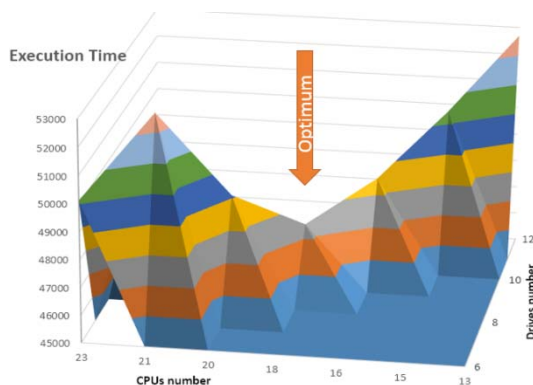
$T_a$  – астрономическое время



При большом количестве процессоров загрузка кластера падает, поскольку процессоры простаивают в ожидании монтирования кассет с данными на драйвы. Следовательно, надо выбирать оптимальное соотношение

2. *Время выполнения пакета заданий в зависимости от количества процессоров и драйвов.*

Стрелкой показан оптимум по числу вычислительных процессоров в кластере и дорогих драйвов. Таким образом, конфигурация, обеспечивающая минимальное время исполнения должна состоять из 18 вычислительных процессоров и 9 драйвов



# 4. Использование параллельных вычислений

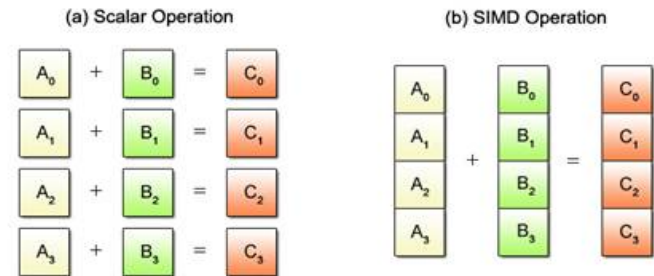
- Введение параллелизма в обработке данных ФВЭ и ЯФ облегчается благодаря естественной организации структуры данных по событиям. Это позволяет легко реализовать многопоточный параллелизм путем одновременной обработки событий на разных процессорах.

- Однако в таких экспериментах, как СВМ, где должны обрабатываться терабайты данных, поступающих за секунду, требуется вводить параллелизм уже на уровне событий. Это можно выполнить, применяя технологию вычислений SIMD (Single Instruction Multiple Data), когда много данных пакуются в один регистр, чтобы выполнять с ними параллельно одну коллективную операцию за цикл процессора.

Сейчас любом современном CPU встроена векторная интеловская SIMD технология **SSE** (Streaming SIMD Extension) с 128 битными регистрами 4 x float. **В будущем:** AVX: 8 x float

Однако переделка C++ кода, использующего скалярные компьютерные команды в векторные команды SSE требует значительной ручной работы по расписыванию циклов в линейную цепочку команд, переделке алгоритмов, отладке и тестированию полученной «симдизированной» программы

**Рассмотрим пример распараллеливания на СВМ в программе А.Лебедева LITtracking**



4 параллельных операции сложения



## Фильтр Калмана для реконструкции треков

Фильтр Калмана(ФК) – это эффективный рекурсивный фильтр, оценивающий состояние линейной динамической системы, используя ряд неточных измерений.

Вектор состояния:  $\vec{x} = (x, y, t_x, t_y, q / p)^T$

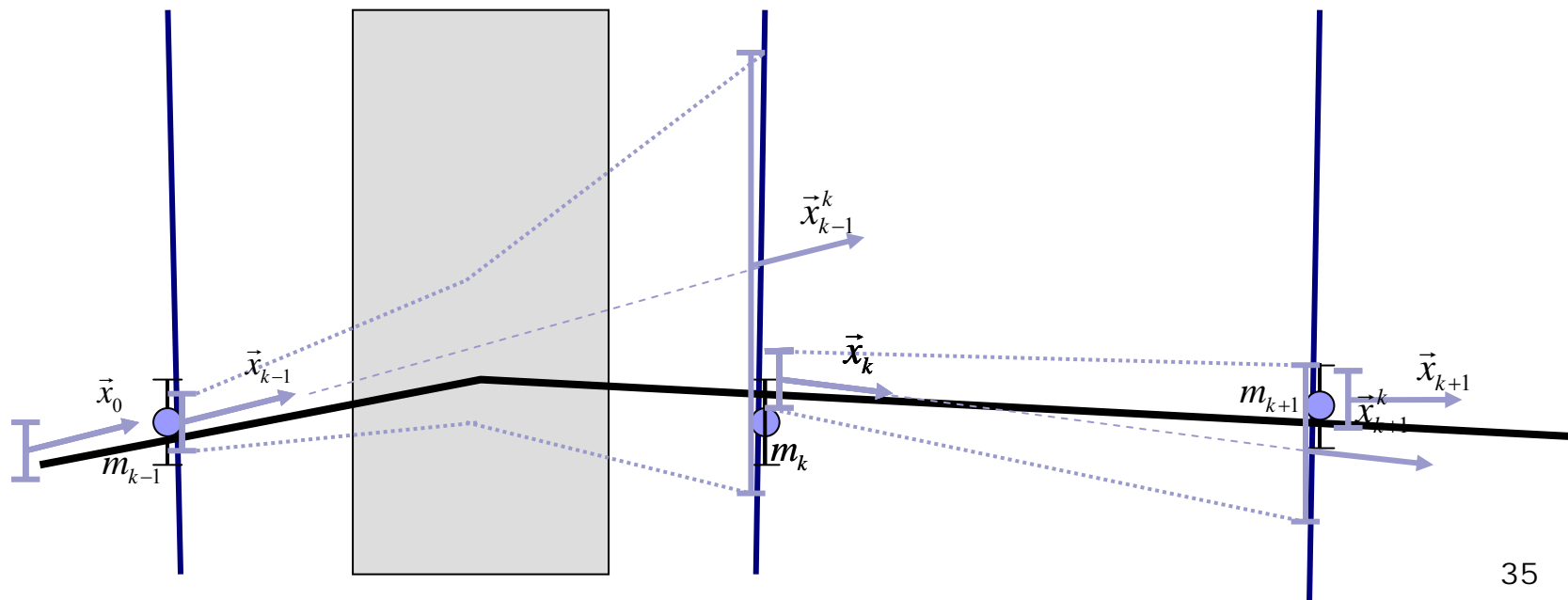
### Преимущества фильтра Калмана (ФК):

- ФК хорошо подходит для одновременного поиска и оценки параметров трека;
- в отличие от МНК необходимо инвертировать только небольшие 5x5 матрицы, что делает ФК быстрым по сравнению с глобальным МНК фитированием;
- позволяет легко учитывать многократное рассеяние и потери энергии

Предварительно необходимо найти начальное значение вектора состояния  $\vec{x}$

Два этапа:

1. Экстраполяция
2. Учет измерения



# Этапы и результаты распараллеливания

## Оптимизация алгоритмов для их эффективной векторизации

1. **Минимизация доступа к основной памяти.** Основное - 3D аппроксимация 70 Мб карты магнитного поля полиномом 5-ой степени в плоскости станции и параболой – между станциями
2. Упрощение геометрии детектора и реализация в ней быстрой навигации
3. Оптимизация вычислений в фильтре Калмана
  - Использование одинарной точности **float** вместо **double**
  - Замена матричных операций в фильтре Калмана на явные вычисления только над ненулевыми элементами матриц
  - Развертка циклов
  - Удаление ветвлений (if then else ..)

## Векторизация алгоритма оценки параметров трека

Так как все треки независимы и экстраполируются одним и тем же алгоритмом, можно экстраполировать параллельно четыре трека после упаковки соответствующих четырех скалярных параметров трека в вектор с помощью SSE инструкций.

## Параллельный алгоритм трекинга

- Параллельные циклы сначала для группы станций детектора потом для каждой станции
- в цикле упаковка 4-х скалярных треков в вектор; векторная экстраполяция треков между станциями и через вещество; добавление отсчета к треку;

**Результаты. Ускорение времени счета** (Компьютер с 2xCPU Intel Core i7 2.67 ГГц, 8 ядер)

1. для алгоритма оценки параметров трека – **2400**,  
Производительность:  **$2 \cdot 10^6$**  треков/с
2. для алгоритма трекинга ~ **500**

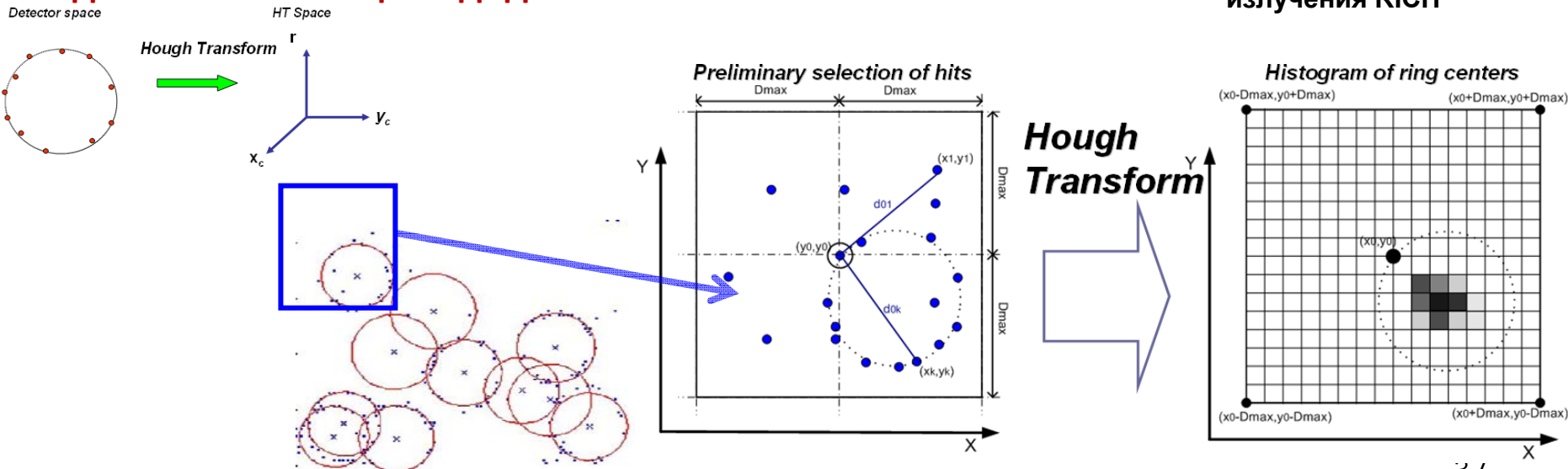
# 5. Совершенствования алгоритмических и программных средств анализа и моделирования

## Пример 2. Детектор RICH

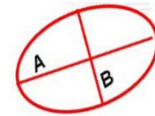
### Этапы обработки (Семен Лебедев)

1. распознать кольца, отсеив ложные
2. учесть оптические искажения детектора, приводящие к эллиптической форме колец;
3. привязать найденные кольца к трекам частиц, интересующих экспериментаторов
4. идентифицировать электроны, подавив пионы

### 1. Быстрое локальное преобразование Хафа для поиска колец-кандидатов



ring-track assignment



→ ring attached to global track

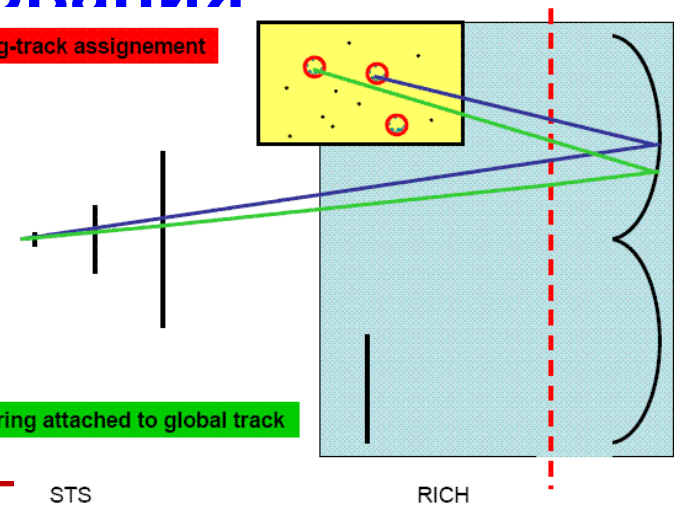


Схема детектора черенковского излучения RICH

# Искусственные нейросети для выбора между гипотезами

Мы здесь не рассматриваем этапы 2-3 подгонки черенковских колец эллипсами и связи их с ближайшими к ним треками, т.к. это - отдельные сложные математические задачи

Нейронные сети применялись для разделения гипотез, чтобы **удалить ложно найденные кольца** и **провести идентификацию частиц**.

Статистическое исследование позволило выбрать 10 наиболее значимых характеристик полученных колец:

1. количество точек в найденном кольце
2. расстояние от центра кольца до ближайшего трека
3. сумма трех наибольших углов между соседними точками
4. число точек в узком коридоре вокруг радиуса
5. радиальная позиция на плоскости фотодетектора
6.  $\chi^2$  эллиптической подгонки кольца
7. большая (A) и (8) малая (B) полуоси эллипса
8. угол поворота эллипса  $\varphi$  относительно оси абсцисс
9. азимутальный угол трека
10. импульс трека

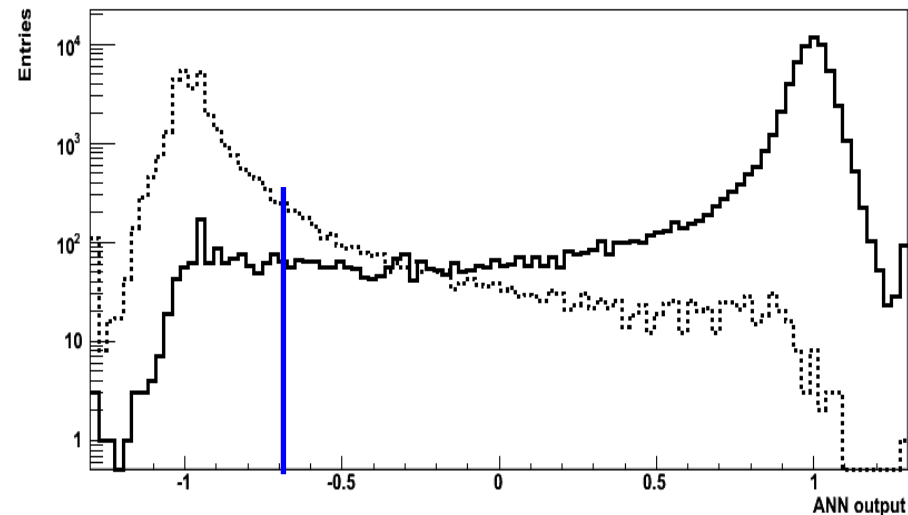
Эти параметры использовались для выброса ложных колец

Обучение нейросети велось на Монте-Карло выборке из событий с 3000 электронов (+1) и 3000 пи-мезонов (-1).

Подобранный порог обеспечил идентификацию электронов с вероятностями ошибок

1-го рода 0.018 и 2-го рода 0.0004

при 90%-й эффективности распознавания колец



Значение выходного нейрона

# Шаги распараллеливания

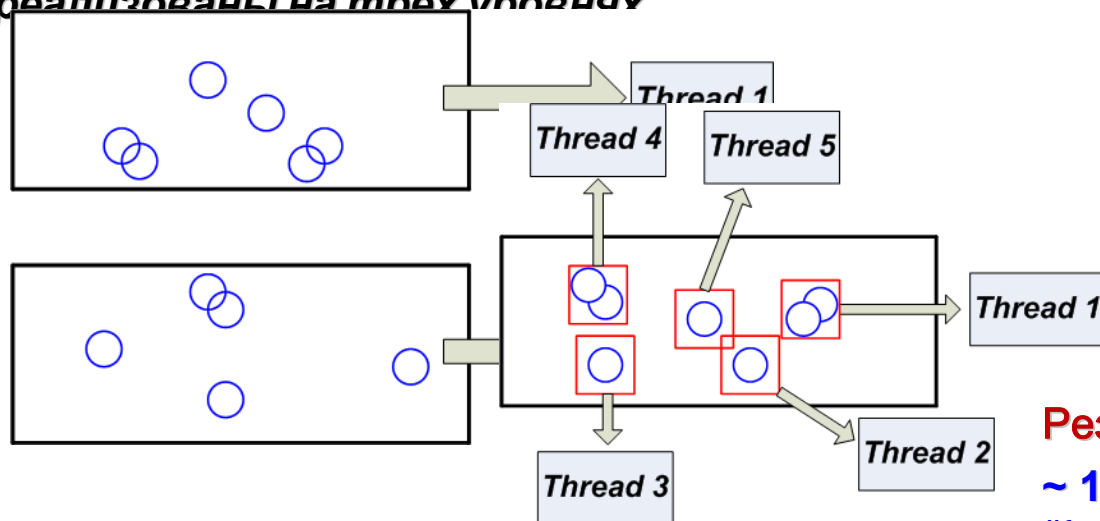
## 1. Векторизация преобразования Хафа

□ Данные RICH детектора были векторизованы в 4-хмерные  
Хит:  $X_v, Y_v$ , где  $X_v=(X_0, X_1, X_2, X_3)$ ;  $Y_v=(Y_0, Y_1, Y_2, Y_3)$

□ Векторизация алгоритма оценки параметров треков  
использовались SSE инструкции вычисления  
параметров в преобразовании Хафа. Оно  
**выполняется одновременно для 4 окружностей!**

## 2. Использование многопоточности в алгоритме

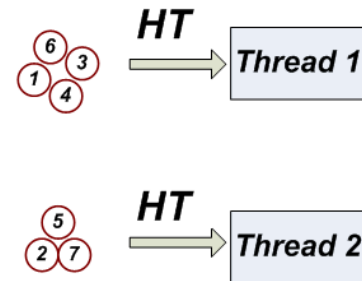
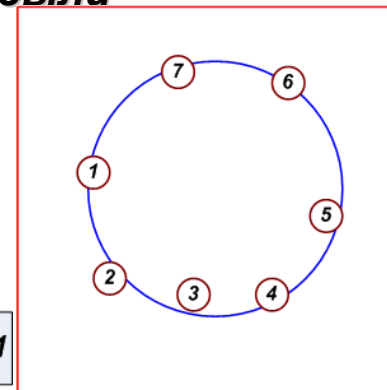
Для алгоритма реконструкции колец параллельные  
вычисления с использованием многопоточности были  
реализованы на трех уровнях



1) Два независимых фотодетектора

2) Параллельный поиск колец  
в разных областях

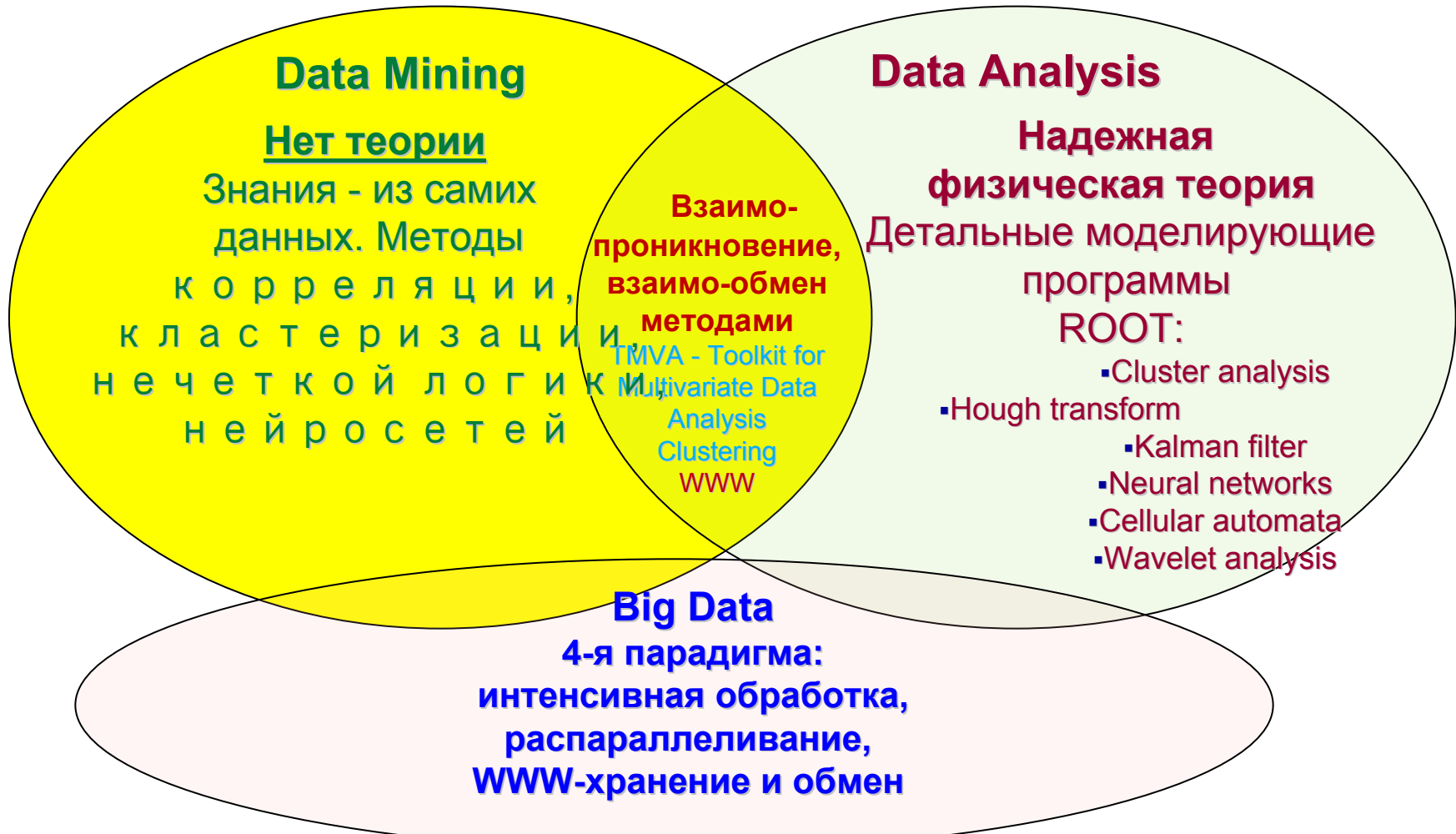
3) Разделение хитов на группы  
и выполнение перебора в  
каждой группе параллельно



**Результаты. Ускорение времени  
~ 150 раз**

(Компьютер 2xCPU Intel Core i7 2.67 ГГц, 8 ядер)





Математика, теория вероятностей и статистика, вычислительные методы, теоретическое программирование




***Спасибо за внимание!***

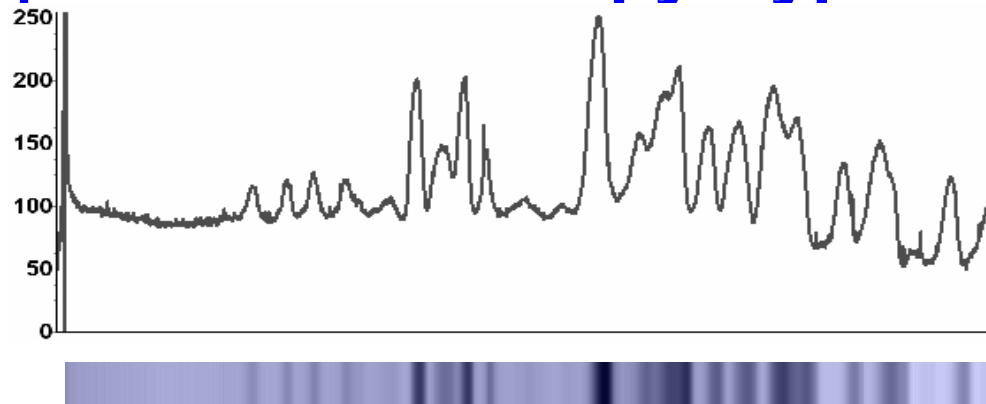


# **Back up slides**

# Example 4. Classification of protein structures

Densitogram, 

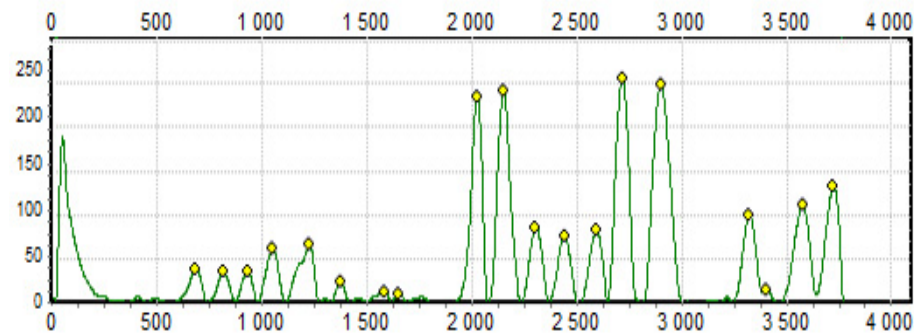
obtained by scanning  
electrophoretic spectrum  
of wheat protein 



**Task: determine the grain variety by its spectrum.**

Preprocessing:

smoothing, subtraction of background, removal of non-linear shifts of lanes relative to each other due to non-stationarity of electrophoretic processes (with the help of a Hopfield neural network), finding significant peaks.



Main problem – large dimensionality of the input vector = 3900 points.

**How to compress it for input into the neural network, preserving information?**

# Curse of Big Data problems

**Input:** 4000 pixels

**Output:** 5 ÷ 20 sorts to be classified

**MLP dimension**  $D=4000*1000+1000*20=$

$4.2 * 10^6$ , i.e **millions of weights**

or equations to solve by the error back propagation method!

**A cardinal reduction of input data was needed**



## Feature extraction approaches

**1<sup>st</sup> approach: Fast Fourier transform (FFT).** Real part of direct FFT was used to transform input data to the frequency domain, where the highest frequencies were cut up to 256 (16 times of reduction)

$$H_n \equiv \sum_{k=0}^{N-1} h_k \exp\left(\frac{kn}{N} 2\pi i\right) \quad n \in \left[-\frac{N}{2}, \frac{N}{2}\right]$$

After transforming all training samples to Fourier space NN-classifier (256/40/5) was trained on them and tested again on transformed sample.

**Result for 5 sorts: 100% of efficiency**

**2<sup>d</sup>. Principal component analysis (PCA)** transforms d-dimensional data to a m-dimensional subspace by using the information of their covariance matrix  $S_X = \text{cov}(X_i X_k)$ . The orthogonal Karhunen–Loeve transform gives the diagonal form of  $S_X$  with eigenvalues  $l_i$  numbered in decreasing order. Thus, we can retain only  $m$  most significant eigenvalues  $l_1, l_2, \dots, l_m$  ( $m \ll p$ ) and express the input data in terms of these principal components as  $X_i \cong l_{1i} Y_1 + l_{2i} Y_2 + \dots + l_{mi} Y_m$ .

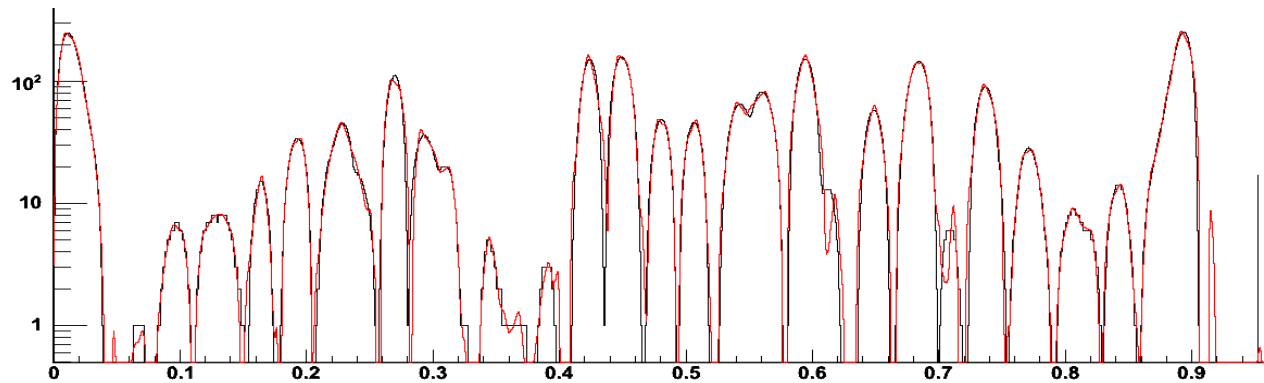
**Result for 5 sorts, m=150: 99,54%**



# Feature extraction approaches (contin)

**3<sup>d</sup>. Discrete wavelet transform (DWT).** Coiflet DWT of the 6<sup>th</sup> order were applied to transform all training and testing samples into wavelet space. Then NN-classifier (256/40/5) was trained and tested on them.

**Result for 5 sorts: 100% of efficiency**



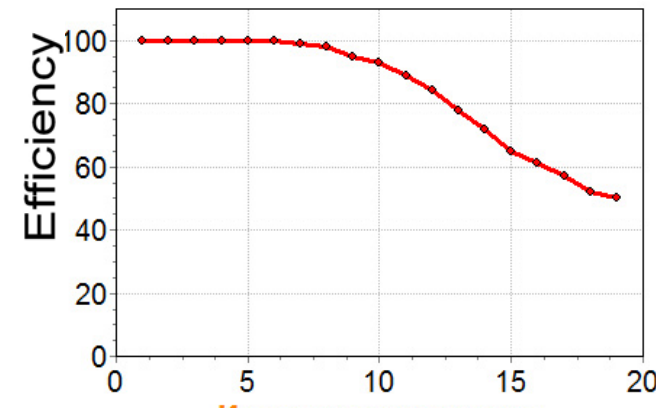
An example of the quality of DWT transform

- original spectrum
- restored spectrum

(ordinata is in logarithmic scale).

## **4<sup>th</sup>. Densitogram vectorization.**

Since peak positions and their amplitudes are the main sort feature of the every densitogram, a special algorithm was developed to attribute unambiguously every peak and its position to the corresponding vector component. It gives data reduction in **two orders of magnitude** and converts each densitogram into **40-dimensional vector**. However ANN efficiency keeps stable while the sorts number is less than 10 only.



Therefore **preclassifying by cluster analysis is now under study**