17.02.2017

Альнаджар Халед Хасан (КНИТУ-КАИ)

ЭФФЕКТИВНЫЙ ГЕНЕРАТОР ПСЕВДОСЛУЧАЙНЫХ ЧИСЕЛ, ОСНОВАННЫЙ НА НЕЧЕТКОЙ ЛОГИКЕ

Разработана и исследована архитектура генератора псевдослучайных чисел, основанного на нечеткой логике (НГПСЧ). Исследуются и выбираются наилучшие параметры НГПСЧ с целью получения последовательностей, удовлетворяющих статистическим критериям качества. Произведена оценка качества сгенерированных псевдослучайных последовательностей с помощью пакета тестов DIEHARD. Произведено сравнение НГПСЧ с другими известными генераторами с помощью наиболее важных тестов набора NIST. Произведен анализ стойкости НГПСЧ к ряду атак.

03.03.2017

В.А. Райхлин, Р.К. Классен (КНИТУ-КАИ)

МОГУТ ЛИ GPU-АКСЕЛЕРАТОРЫ СУЩЕСТВЕННО ПОВЫСИТЬ ЭФФЕКТИВНОСТЬ КОНСЕРВАТИВНЫХ СУБД ЗНАЧИТЕЛЬНЫХ ОБЪЕМОВ НА КЛАСТЕРНОЙ ПЛАТФОРМЕ?

Обсуждаются вопросы построения СУБД консервативного типа (с эпизодическим обновлением данных в специально выделяемое время) на платформе GPU-кластеров при объемах баз данных – $V_{БД}$ не менее 100GB. Их актуальность определяется современными тенденциями интеллектуальной обработки больших информационных массивов с применением графических ускорителей – GPU. По условию обработка запросов ведется по регулярному плану. В узлах кластера под управлением СУБД MySQL функционируют многоядерные процессоры. В динамике обработки запросов узловая БД оказывается в оперативной памяти узла объемом до 128 GB. Рассматриваются случаи средних $V_{БД}$ – вблизи 100GB, реплицируемых по узлам, и достаточно больших $V_{БД}$ – от сотен GB до единиц TB, хешируемых на множестве узлов. В первом случае анализируются два варианта функционирования СУБД: 1) на CPU – операции *«select-project»*, на GPU – *«join»*; 2) на CPU – *«project»* и *«join»,* на GPU – *«select»*, БД хранится в сжатом виде. Установлено, что оба варианта использования ускорителей неконкурентоспособны. Во втором случае хешируется сжатая БД по узлам IO с ускорителями, на которых выполняются операции *«select-project»*, операции *«join»* реализуются на узлах JOIN без GPU. Приведено теоретическое обоснование такой организации. Экспериментальное подтверждение значительного превышения ее эффективности по сравнению с ранее разработанной СУБД *Clusterix-M* без GPU связывается с разработкой натурной модели *Clusterix-G*.

17.03.2017

В.М. Захаров, С.В. Шалагин, Б.Ф. Эминов (КНИТУ-КАИ)

АВТОМАТНАЯ МОДЕЛЬ ПРЕДСТАВЛЕНИЯ НЕЛИНЕЙНЫХ МАКСИМАЛЬНЫХ ПСЕВДОСЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ НАД КОНЕЧНЫМ ПОЛЕМ

Представлена автоматная модель построения нелинейных псевдослучайных последовательностей с периодом $L = 2^n$. Функция переходов автомата реализуется на основе генератора *M*-последовательности. Функция выхода определяется как нелинейная функция усложнения, реализующая векторное однозначное преобразование состояний автомата на основе системы нелинейных модулярных операций по модулю, принадлежащему к множеству простых чисел Ферма. Размер ансамбля выходных последовательностей является экспоненциальной функцией от величины *n*.

14.04.2017

Я. А. Фурман, И. Л. Егошина, К.О. Иванов (ПГТУ, г. Йошкар-Ола)

МЕТОД И АЛГОРИТМЫ ЛОКАЛЬНОГО АНАЛИЗА ЭЛЕКТРОЭНЦЕФАЛОГРАММ НА БАЗЕ ИХ КОНТУРНЫХ МОДЕЛЕЙ

Методы количественной обработки данных в клинической электроэнцефалографии в настоящее время базируются на спектральных и корреляционных видах анализа. Как следует из механизмов их действия, основные результаты компьютерной обработки носят интегральный характер. Они получаются при усреднении с теми или иными весами по эпохам всех отсчетов анализируемой ЭЭГ. В случае, если анализируемая эпоха ЭЭГ содержит набор разнородных паттернов, то часть информации о разнородных колебаниях в результате усреднения утрачивается. С целью обеспечения идентификации разнородных паттернов в ЭЭГ-сигнале в автоматическом режиме в работе предложен подход к обработке ЭЭГ-сигнала, заключающийся в его сегментации и анализе каждого отдельного колебания. Разработана математическая модель сигнала ЭЭГ в виде вектора унитарного пространства – контурная модель ЭЭГ. Модель допускает декомпозицию сигнала на отдельные информативные фрагменты и получение количественных характеристик их форм на основе математического аппарата контурного анализа. Определены информативные признаки волн ЭЭГ, достаточные для их классификации. Разработаны алгоритмы вычисления информативных характеристик. Предложен метод классификации электроэнцефалограммы, включающий процедуру ее сегментации, определение информативных признаков сегментов, сравнение информативных признаков с диапазонами значений, принятыми в клинической практике, классификацию всей электроэнцефалограммы по совокупности классов составляющих ее волн. Метод позволяет учитывать проявление единичных патологических комплексов и выявлять пограничный характер электроэнцефалограмм. Получены положительные результаты классификации пограничных состояний на реальных электроэнцефалограммах, классифицируемых как «норма» при помощи классических методов анализа.

28.04.2017

В.Н. Поляков (Institute of Linguistics of Russian Academy of Sciences)

НЕПОЛНЫЙ СИНТАКСИЧЕСКИЙ РАЗБОР В МОДЕЛИ ЗАВИСИМОСТЕЙ И ЗАДАЧИ НА ЕГО ОСНОВЕ

В докладе рассматривается новое направление обработки естественно-языкового текста (ОЕЯТ), основанное на неполном синтаксическом разборе. Работы в этом направлении ведутся с 2012 года в рамках НИТУ «МИСиС» в сотрудничестве с КФУ. За прошедшие 5 лет создана библиотека NLP@Cloud, которая позволяет производить токенизацию, морфологический анализ, частичный синтаксический анализ (чанкинг), орфокоррекцию, анализ би-грамм для русского языка и аналогичные этапы ОЕЯТ для английского языка. Библиотека построена на основе фреймворка UIMA с использованием языка программирования Java. Неполный синтаксический разбор (чанкинг) строится на базе синтаксической модели Теньера. При этом в модель синтаксического анализа внесен ряд эвристик, представляющих научную новизну и открывающих новые возможности ОЕЯТ. В частности, неполный синтаксический разбор позволяет сразу перейти к синтаксической модели предложения и далее – к семантике, не дожидаясь полного описания грамматики языка. Модели зависимостей легче трансформируются в логическую нотацию или нотацию фреймов, чем модель непосредственных составляющих (грамматики Хомского). В целом, реализация чанкинга для русского и английского языков – это еще один шаг к пониманию текста. Следующим шагом будет создание семантических моделей (сентимент-анализ, голосовые помощники, естественно-языковой интерфейс, извлечение знаний и т.д.) с использованием неполного семантического анализа.

12.05.2017

Р.Ф. Гибадуллин, А. Г. Савельев (КНИТУ-КАИ)

ПРИНЦИПЫ ИСПОЛЬЗОВАНИЯ РАЗДЕЛЯЕМОЙ ПАМЯТИ GPU NVIDIA ПРИ ВЫПОЛНЕНИИ ЗАПРОСОВ К КАРТОГРАФИЧЕСКИМ БАЗАМ ДАННЫХ

В докладе представлены результаты исследования роли конфликтов банков разделяемой памяти при выполнении запросов к картографическим базам данных в виде OLAP-структур на графических процессорах NVIDIA. Формулируются предложения по оптимизации банков памяти для повышения скорости обработки запросов к таким БД.

26.05.2017

И.С. Вершинин (КНИТУ-КАИ)

РЕЛЕВАНТНЫЙ МЕТОД АНАЛИЗА СТОЙКОСТИ АССОЦИАТИВНОЙ СТЕГОЗАЩИТЫ

На основании проведенных исследований формулируется стеганографический метод анализа стойкости ассоциативной защиты. Педлагаемый метод учитывает специфику ассоциативной стеганографии. Приводится обоснование метода на множестве конкретных атак.