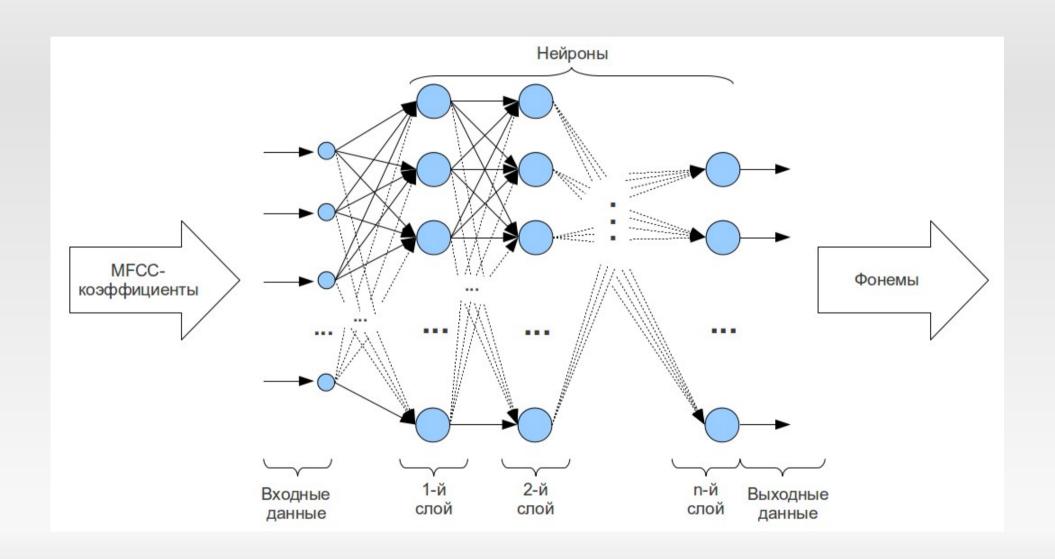
Разработка и исследование нейросетевых алгоритмов распознавания усной речи для параллельных архитектур CUDA и OpenMP

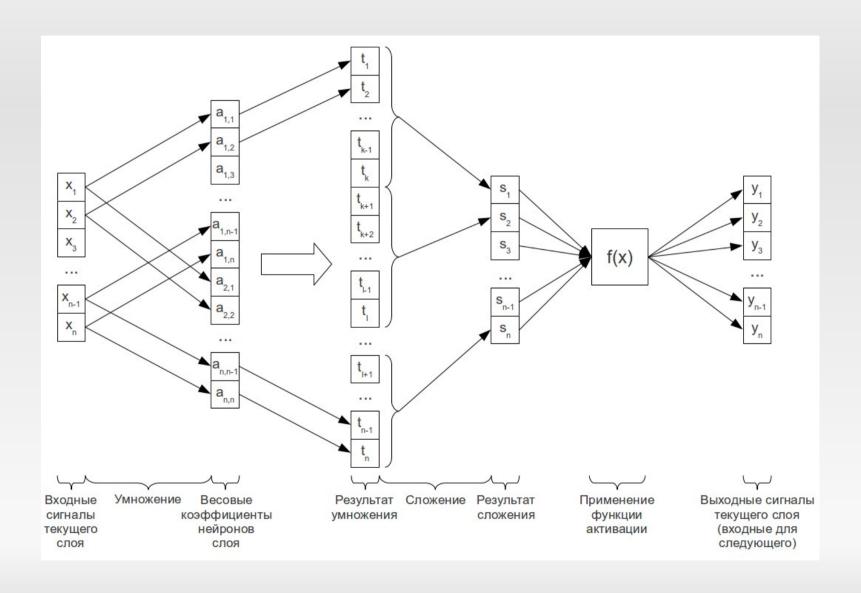
Автор: Шатохин Н.А.

Руководитель: Бондаренко И.Ю.

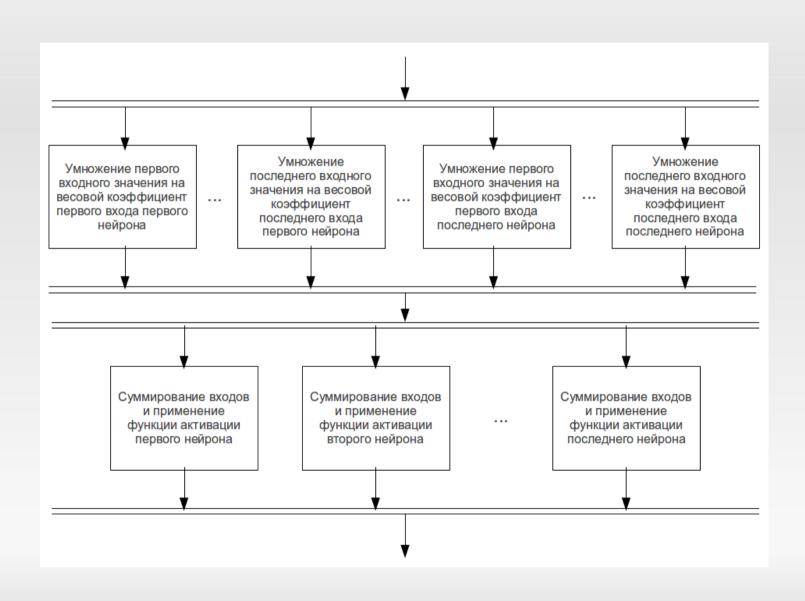
Нейронная сеть для пофонемного распознавания речи



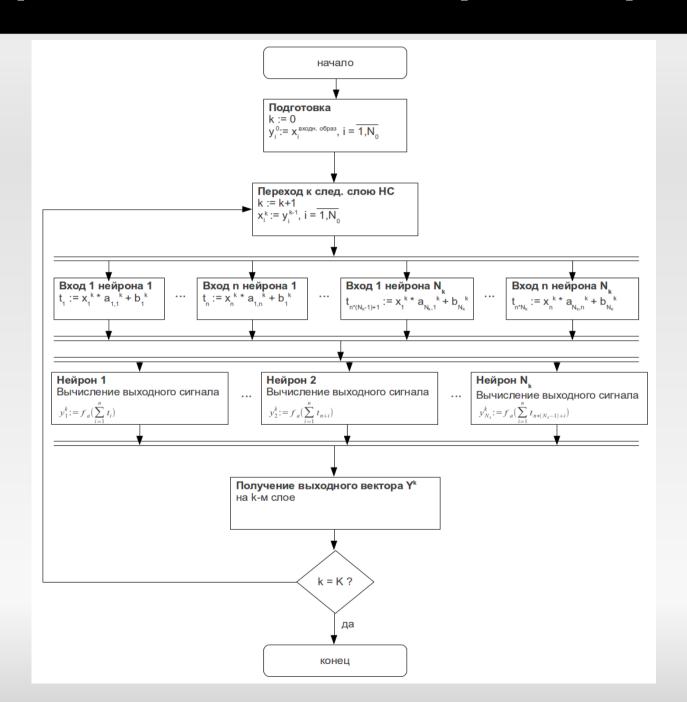
Моделирование нейронной сети



Распараллеливание нейроалгоритма



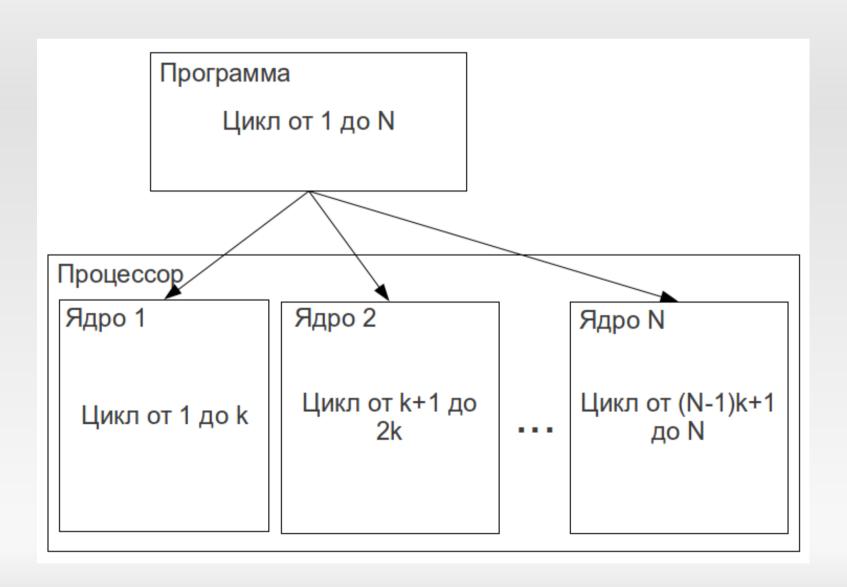
Распараллеливание нейроалгоритма



Технологии распараллеливания

 OpenMP (Open Multi-Processing) — открытый стандарт для распараллеливания программ на языках Си, Си++ и Фортран. Описывает совокупность директив компилятора, библиотечных процедур и переменных окружения, которые предназначены для программирования многопоточных приложений на многопроцессорных системах с общей памятью.

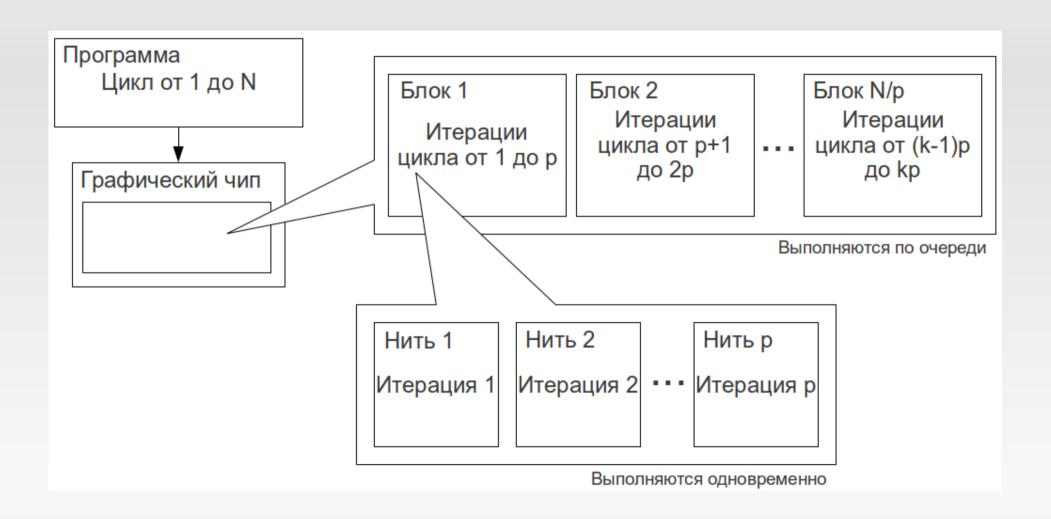
Как работает OpenMP



Технологии распараллеливания

• CUDA (англ. Compute Unified Device Architecture) — программно-аппаратная архитектура, позволяющая производить вычисления с использованием графических процессоров NVIDIA, поддерживающих технологию GPGPU (произвольных вычислений на видеокартах).

Как работает CUDA



Тестирование

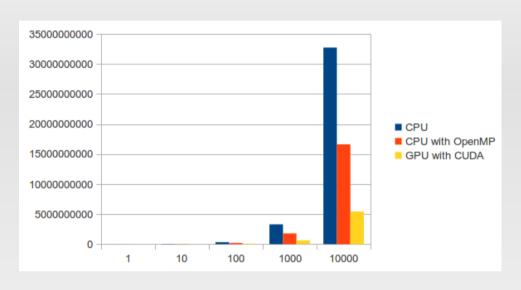
Тестирование проводилось на компьютере со следующими характеристиками:

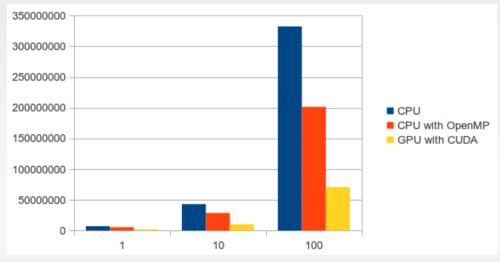
Процессор: Intel Core2Duo T7500 2,2GHz, 2 ядра

Видеокарта: Nvidia GeForce 8600M GS, 800MHz, 32 потоковых процессора, 512 нитей в блоке

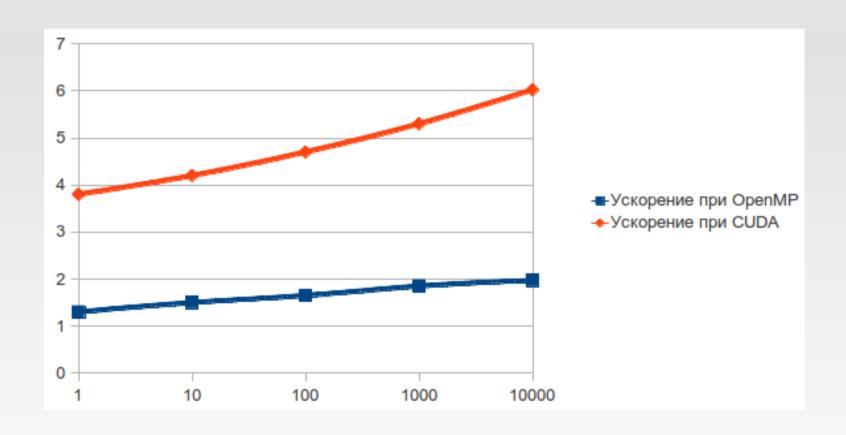
Для тестирования была создана четырехслойная сеть с 400 входами и 454, 300, 150, 41 нейронами в каждом слое соответственно. Были обработаны 5 файлов, содержащих 1 набор входных сигналов, 10, 100, 1000 и 10000 соответственно.

Результаты

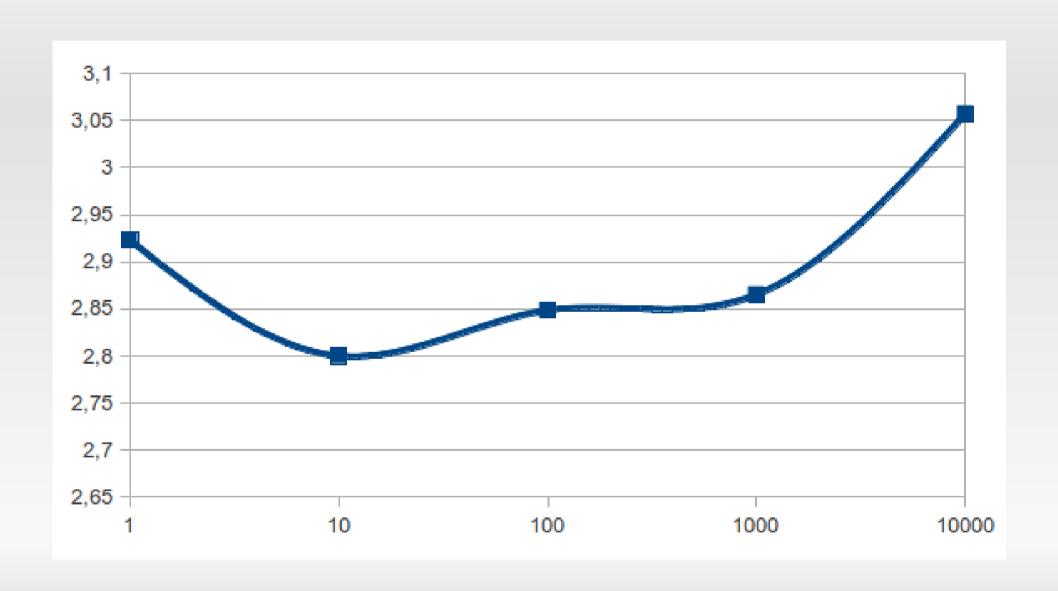




Ускорение обработки



Ускорение при использовании CUDA по сравнению с OpenMP



Спасибо за внимание!