

Летняя международная XXXI молодёжная Школа-конференция по
параллельному программированию

Разработка параллельного алгоритма для задачи распространения близости

Исполнитель проекта: Чикибаева Д.Ю.

Студентка 3 курса, факультета информационных технологий КазНУ им.
Аль-Фараби

Руководитель проекта: Мансурова М.Е.

Дата доклада: 13.07.2018

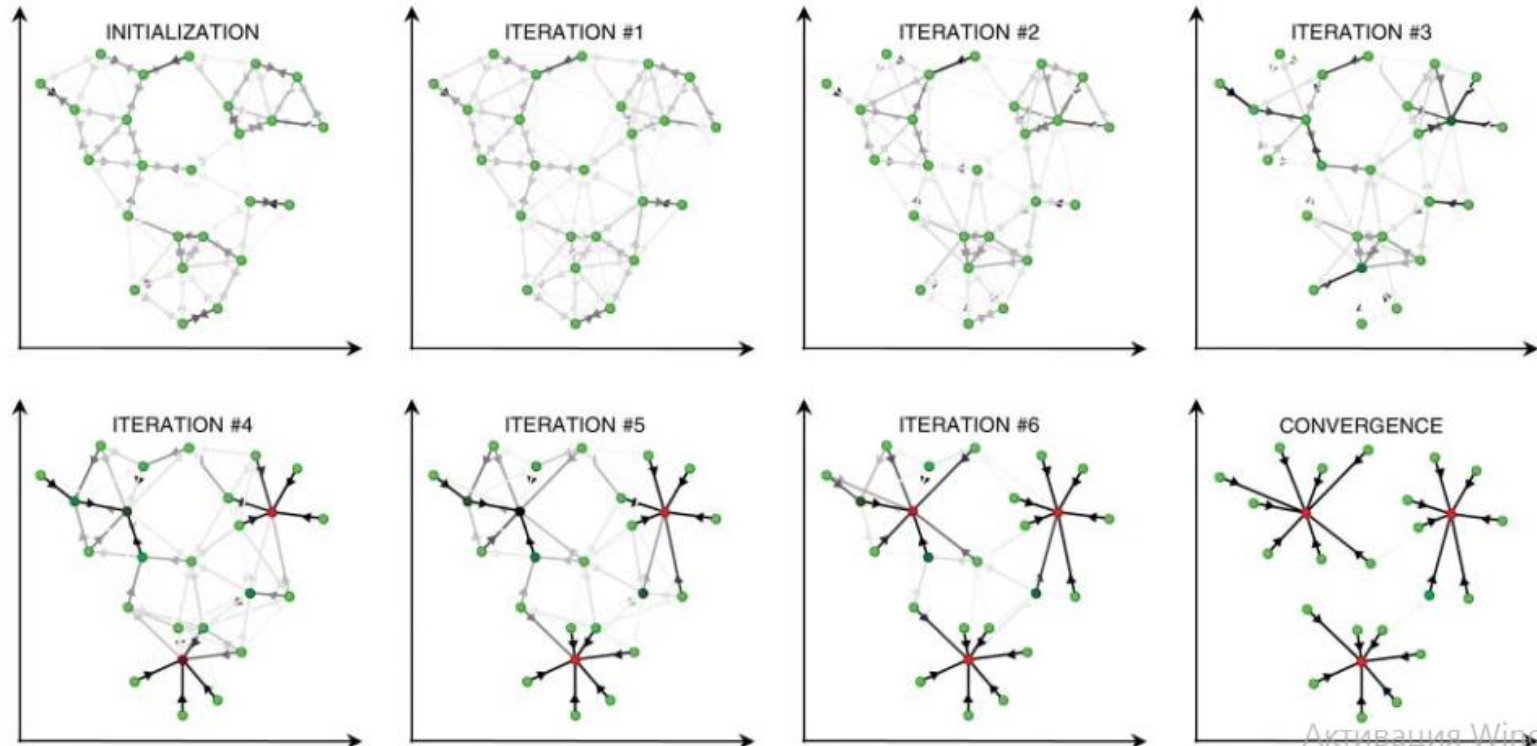
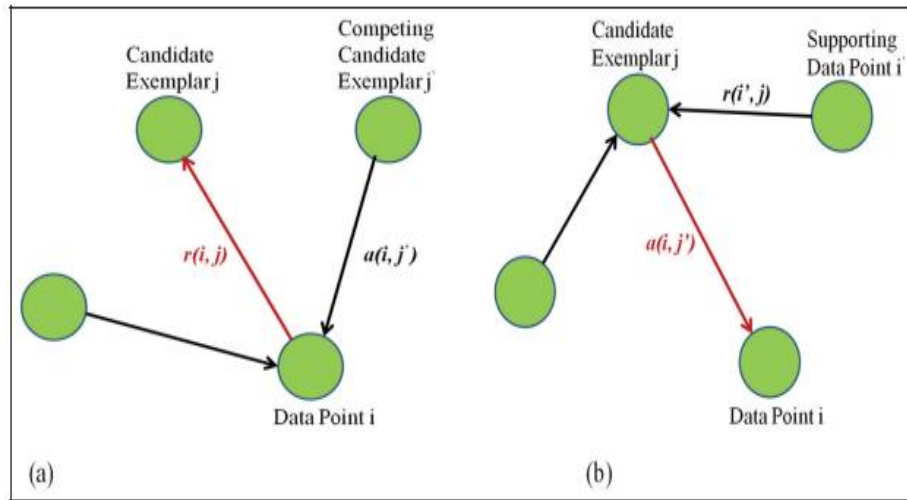
План доклада

- 1. Постановка задачи
- 2. Идея решения
- 3. Реализация
- 4. Тестирование
- 5. Заключение

Постановка задачи

- Кластеризация. Разбиение на кластеры данных с плавающей точкой с использованием метода распределения близости.
- Параллелизация вычислений алгоритма для уменьшения времени выполнения с использованием директив openMP, MPI.

Реализация



$$r(i, j) = s(i, j) - \max_{j' \neq j} \{a(i, j') + s(i, j')\}$$

$$a(i, j) = \min \left\{ 0, r(j, j) + \sum_{i' \neq i, j} \max\{0, r(i', j)\} \right\}, i \neq j$$

$$c_j \leftarrow \arg \max_{1 \leq j \leq N} [r(i, j) + a(i, j)]$$

Результат

- Алгоритм был использован с различным количеством входных данных.
- 100 точек – 1мин 8сек
- 1000 точек – 20мин 14сек

Заключение

- Результаты:

Создана система с получением данных, объединённых в кластеры

- Направление развития проекта:

Использование алгоритма для построения концепт-карт из неструктурированного текста для извлечения и представления знаний.