ИПМ им.М.В.Келдыша РАН

Абрау-2016 • Труды конференции

В.А. Серебряков

**Обзор алгоритмов
локально-чувствительного
хеширования**

Рекомендуемая форма библиографической ссылки

Серебряков В.А. Обзор алгоритмов локально-чувствительного хеширования // Научный сервис в сети Интернет: труды XVIII Всероссийской научной конференции (19-24 сентября 2016 г., г. Новороссийск). — М.: ИПМ им. М.В.Келдыша, 2016. — С. 303-303. — doi:[10.20948/abrau-2016-8](https://doi.org/10.20948/abrau-2016-8)

Размещена также [презентация к докладу](#)

Обзор алгоритмов локально-чувствительного хеширования

В.А. Серебряков

ВЦ РАН

Аннотация. Часто разработчики информационных систем сталкиваются с задачей попарного сравнения записей, например, библиографических записей, веб страниц и т.д. Эта задача сформулирована давно и заключается она в следующем: даны две записи. Спрашивается: насколько они отличаются при определенном способе вычисления этого отличия? Задача хорошо проработана и существует много подходов для ее решения. Все они основаны на попарном сравнении записей. Однако, если записей очень много (миллионы), попарное сравнение невозможно. Нужно привлекать алгоритмы, которые, во-первых, выделяют из всего множества записей априорно близкие и, во-вторых, сводят сравнение самих записей к сравнению их простых образов.