

Мишота А.А., Ковалева И.Л.

Белорусский национальный технический университет

Как правило, стандартный функционал современных операционных систем предоставляет пользователю возможность поиска требуемой программы или файла по ключевым словам в их названии. При этом большинство специализированных поисковых утилит являются платными. Поэтому разработка приложения, обеспечивающего поиск требуемого контекста по ключевым словам, находящимся не только в названии, но и в самом файле, остается актуальной задачей.

В качестве критерия оценки работы приложения предлагается использовать «время поиска». Параметрами оптимизации являются алгоритм поиска, размеры буферов чтения и записи (в байтах) и количество потоков, ограничениями являются: размеры буферов от 1 до n байт (где n – объём доступной оперативной памяти компьютера), количество потоков от 1 до $2m$ (где m – количество логических процессоров на компьютере).

Для поиска заданного контекста в наборе данных в приложении может быть использован один из следующих алгоритмов поиска: алгоритм поиска Бойера - Мура, алгоритм Рабина - Карпа, алгоритм Кнута-Морриса-Пратта, алгоритм Чжу - Такаоки и алгоритм «Быстрый поиск».

Алгоритм Бойера – Мура является одним из самых быстрых алгоритмов. Следует, однако, заметить, что его эффективность растет лишь с увеличением длины строки и, соответственно, длины образца.

Так, при длине строки меньшей или равной 10 символов, он показал себя хуже, чем последовательный поиск.

Аналогичные результаты показывает и алгоритм Кнута-Морриса-Пратта, как для коротких, так и для длинных слов.

Его можно использовать как универсальный, когда неизвестны длины строки и образца. Алгоритм Рабина-Карпа, при его схожести с последовательным поиском, работает быстрее, а его простота и малые трудозатраты на реализацию, делают его привлекательным для использования в специальных программах.

Лучший результат показал алгоритм быстрого поиска, который является улучшенной реализацией алгоритма поиска Бойера - Мура.

Приложение разработано на языке C#. В нем реализованы все указанные выше алгоритмы. При этом алгоритм Бойера - Мура используется в приложении «по умолчанию».