97. ПРИМЕНЕНИЕ АЛГОРИТМА БОЙЕРА–МУРА ДЛЯ РЕШЕНИЯ ЭКОНОМИЧЕСКИХ ЗАДАЧ

Жданович А.Д., студентка гр. 573904

Белорусский государственный университет информатики и радиоэлектроники г. Минск, Республика Беларусь

Кириенко Н.А. – канд. техн. наук, доцент каф. ЭИ

Аннотация: в работе рассматриваются особенности алгоритма Бойера–Мура и его модификаций в контексте решения экономических задач. Показано, что алгоритм обеспечивает высокую эффективность поиска в больших массивах текстовых данных за счёт использования эвристик смещения. Приведены основные области применения алгоритма в экономике и проанализированы его преимущества по сравнению с традиционными методами поиска.

В условиях цифровизации экономики существенно возрастает объём обрабатываемой информации [1]**.** Финансовые системы, электронная коммерция и корпоративные информационные системы ежедневно генерируют большие массивы текстовых данных: логи транзакций, отчёты, клиентские отзывы и другие документы. Это требует использования эффективных алгоритмов поиска и анализа информации.

Одним из наиболее эффективных алгоритмов поиска подстроки является алгоритм Бойера–Мура [2]**.** Его особенность заключается в сравнении символов шаблона с текстом справа налево, а также использовании эвристик «плохого символа» и «хорошего суффикса». Благодаря этому алгоритм способен пропускать значительные участки текста, что существенно снижает количество операций. В среднем его временная сложность приближается к O(n).

Алгоритм Бойера–Мура широко применяется в экономических задачах. Одним из основных направлений является анализ логов транзакций и финансовый мониторинг [1]. С его помощью можно быстро выявлять подозрительные операции, находить ключевые коды и анализировать большие потоки данных, что особенно важно для систем обнаружения мошенничества.

Другой важной областью применения является обработка текстовых данных клиентов. Компании анализируют отзывы, обращения и сообщения пользователей для повышения качества обслуживания. Алгоритм позволяет эффективно находить ключевые фразы и ускоряет обработку больших массивов информации.

Также алгоритм применяется при работе с базами данных и электронным документооборотом. Он позволяет ускорить поиск подстрок в текстовых полях, что важно при анализе отчётов, договоров и других документов. В системах электронной коммерции алгоритм используется для обработки поисковых запросов пользователей и фильтрации иинформации [4].

Помимо классического алгоритма Бойера–Мура существуют его модификации, такие как Boyer–Moore–Horspool [3] и алгоритм Sunday. Эти варианты используют упрощённые эвристики и часто показывают более высокую производительность на практике. Для коротких шаблонов могут применяться бит-параллельные алгоритмы.

В рамках исследования было проведено сравнение производительности классического алгоритма Бойера–Мура, его модификации Horspool и прямого (default) алгоритма поиска. Эксперименты выполнялись на примере поиска шаблонов в текстовом массиве большого размера. Рассматривались два сценария: поиск шаблона, расположенного в конце текста, и поиск шаблона из середины текста.

Результаты экспериментов для первого случая представлены на рисунке 1. Как видно из полученных данных (см. рис. 1), алгоритм Horspool показал наименьшее время выполнения (около 18 мс), классический алгоритм Бойера–Мура также продемонстрировал высокую эффективность (около 25 мс), тогда как прямой алгоритм поиска оказался значительно медленнее (около 120 мс). Это объясняется тем, что при работе с длинными шаблонами алгоритм выполняет значительные сдвиги и пропускает большие участки текста.

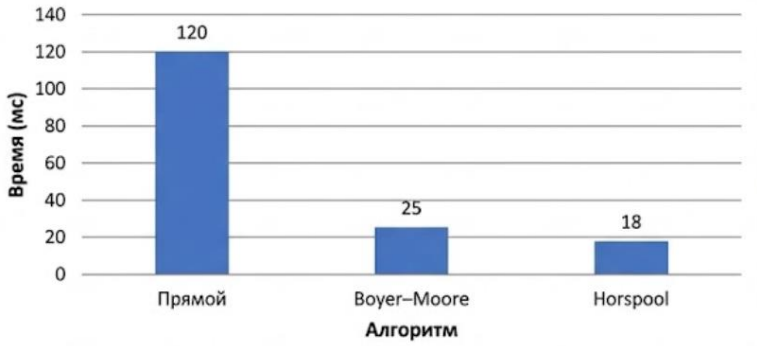

Рисунок 1 – Сравнение времени поиска шаблона различными алгоритмами (шаблон расположен в конце текста)

Аналогичные результаты получены и для второго сценария, где шаблон расположен в середине текстового массива (см. рис. 2). В данном случае алгоритм Horspool также показал наилучший результат (около 15 мс), алгоритм Бойера–Мура — около 20 мс, тогда как прямой алгоритм выполнялся значительно дольше (около 80 мс). Несмотря на изменение абсолютных значений, общее соотношение эффективности алгоритмов сохраняется.

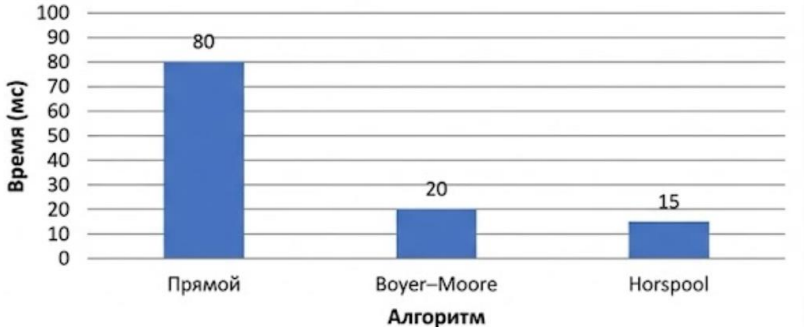

Рисунок 2 – Сравнение времени поиска шаблона различными алгоритмами (шаблон расположен в середине текста)

Дополнительно было установлено, что в тестах типа QuickBench модификация Boyer–Moore–Horspool показала наилучшие результаты [4]. Время выполнения поиска оказалось в несколько раз меньше по сравнению со стандартной функцией поиска строк. Это подтверждает эффективность использования алгоритмов семейства Бойера–Мура при работе с большими текстами и сложными шаблонами.

Несмотря на значительные преимущества, алгоритм имеет и ограничения. Он менее эффективен при поиске очень коротких шаблонов или при необходимости одновременного поиска большого числа различных образцов. В таких случаях используются альтернативные методы поиска.

В заключение можно отметить, что алгоритм Бойера–Мура и его модификации являются эффективным инструментом обработки текстовых данных в экономике. Их применение позволяет существенно ускорить поиск информации, снизить вычислительные затраты и повысить эффективность работы информационных систем. В условиях роста объёмов данных значение подобных алгоритмов будет только увеличиваться.

Список использованных источников:

1. *Алгоритм Бойера-Мура в глубоком анализе трафика: как работает и зачем нужен [Электронный ресурс]. – Режим доступа: https://www.securitylab.ru/blog/personal/Technolady/355642.php, свободный. – Дата доступа: 23.03.2026.*
2. *Boyer–Moore string-search algorithm [Электронный ресурс]. – Режим доступа: https://en.wikipedia.org/wiki/Boyer%E2%80%93Moore_string-search_algorithm, свободный. – Дата доступа: 23.03.2026.*
3. *Boyer–Moore–Horspool algorithm [Электронный ресурс]. – Режим доступа: https://en.wikipedia.org/wiki/Boyer%E2%80%93Moore%E2%80%93Horspool_algorithm свободный. – Дата доступа: 23.03.2026.*
4. *Bartłomiej Filipek. Speeding up Pattern Searches with Boyer-Moore Algorithm from C++17 [Электронный ресурс]. – Режим доступа: https://www.cppstories.com/2018/08/searchers/, свободный. – Дата доступа: 23.03.2026.*