

МЕТОД И ОДНОРОДНОЕ УСТРОЙСТВО МАТРИЧНОГО ПОИСКА ПРОСТОГО И СОСТАВНОГО ОБРАЗЦОВ

DOI: 10.36724/2072-8735-2021-15-8-44-51

Титенко Евгений Анатольевич,
Юго-Западный государственный университет,
Курск, Россия, johntit@mail.ru

Manuscript received 25 February 2021;
Accepted 20 April 2021

Щитов Алексей Николаевич,
Юго-Западный государственный университет,
Курск, Россия, a.n.schitov@mail.ru

Keywords: характеристическая матрица,
диагональный поиск, вхождение, строковый операнд

Исследование состоит в развитии аппаратных методов поиска по образцу, которые используют принципы коллективного доступа и параллельной обработки данных, свойственные ассоциативной памяти. Сокращение времени поиска достигается за счет формирования и параллельной обработки двоичной (характеристической) матрицы сравнений символов образца и текста. Введен составной образец, он позволяет гибко описывать условия поиска. Характеристическая матрица имеет форму параллелограмма, она состоит из сдвинутых вправо строк, начиная с первой строки. Такая форма позволяет вести параллельный поиск простого и составного образцов по диагональным элементам матрицы. Разработанный метод аппаратно поддерживает поиск в характеристической матрице. Он отличается одновременным учетом локальных и распределенных связей между элементами диагоналей и строк матрицы. Раздельное вычисление стартовых значений по диагоналям матрицы позволяет независимо вести поиск двух типов образцов за счет вычисления в ячейках диагоналей характеристической матрицы двух выходных поисковых функций. Метод имеет линейную временную и квадратичную аппаратную сложности. Длительность тактового импульса определяется суммой задержек схемы сравнения пары символов, D-триггера и двухходового элемента И. Однородная структура устройства матричного поиска, использование типовых операций позволяют реализовывать устройство на перспективной элементной базе ПЛИС, что определяет его применение в высокопроизводительных системах обработки и передачи разнородной информации.

Информация об авторах:

Титенко Евгений Анатольевич, доцент кафедры ПИ, доцент, к.т.н., Юго-Западный государственный университет, Курск, Россия

Щитов Алексей Николаевич, младший научный сотрудник ЦПИР, аспирант, Юго-Западный государственный университет, Курск, Россия

Для цитирования:

Титенко Е.А., Щитов А.Н. Метод и однородное устройство матричного поиска простого и составного образцов // Т-Comm: Телекоммуникации и транспорт. 2021. Том 15. №8. С. 44-51.

For citation:

Titenko E.A., Schitov A.N. (2021) The method and multiple device of a matrix search for simple and complex samples. *T-Comm*, vol. 15, no.8, pp. 44-51. (in Russian)

Введение

Ассоциативная память применяется в вычислительных машинах, интеллектуальном сетевом оборудовании, распределенных информационно-управляющих системах для построения высокопроизводительных систем обработки и передачи разнородной информации [1-3]. Сетевые информационно-поисковые системы, экспертные динамические системы, автоматизированные системы управления подвижными техническими объектами, системы навигации и сопровождения воздушных космических аппаратов и судов, аппаратно-программные комплексы обработки цифровых потоков (вхождение в синхронизм, поиск маркеров и т.д.) используют ассоциативную память в составе специализированных устройств-акселераторов поиска характеристик, признаков объектов или внешней среды [4-6].

В зависимости от особенностей предметной области задача поиска имеет множество формализаций, направленных на описание процессов поиска и ограничений. Среди них выделяется группа методов и алгоритмов поиска по образцу, использующих теории алгоритмов, теории информации и кодирования для задания поисковых действий с применением шаблонов, эталонных описаний объектов, состояний внешней среды [7,8].

Несмотря на огромное количество подходов и их реализаций (программных, аппаратных), временная сложность методов и алгоритмов поиска по образцу в большинстве случаев является неудовлетворительной вследствие переборного характера вычислений [9]. Задачи on-line поиска (перехвата) в телекоммуникационных, радиотехнических системах, распознавания нештатных состояний на техногенно опасных объектах, мониторинга критических ситуаций (отказов) в системах безопасности и жизнеобеспечения, обработки аэрокосмических изображений высокого разрешения, обеспечения параллельных вычислений в многопроцессорных системах, поддержки и структуризации Big Data, аппаратной поддержки сетевой трейдерской деятельности требуют отклика на динамично меняющуюся обстановку. Как правило, оперативный отклик сводится к линейной временной зависимости поиска по образцу от размера операндов. Особенно остро время поиска проявляется в задачах большой размерности, для которых пространство перебора вариантов оценивается десятками тысяч состояний [10].

Проблемная ситуация заключается в том, что современные программные реализации, использующие циклические конструкции для организации перебора, преимущественно имеют квадратичную временную сложность, что не позволяет их использовать для проблемно-поисковых задач, задач оперативного управления сложными объектами в динамично меняющейся обстановке, интеллектуальных задач параллельной обработки символьной информации большой размерности. Обработка частичных вхождений образца в обрабатываемом тексте связана с необходимостью выполнять множественные отступы (возвраты) как в позициях образца, так и в позициях текста. Эти множественные отступы приводят к непродуктивным затратам времени на выполнение избыточных сравнений.

Известные методы выполнения строковых сравнений (метод прямого сопоставления, методы с предобработкой, с индексированием префиксов строк, с управляемым шагом

сдвига, метод хеширования данных, конвейерный метод и др.) частично или полностью наследуют проблему множественных отступов при поиске полного вхождения образца в текст [9,11].

Разрешение данной проблемной ситуации связывается с созданием аппаратно-ориентированного матричного метода, основанного на операциях множественного сравнения, выбора элементов, параллельного вычисления позиций переходов, имеющих схемотехническую реализацию.

Постановка задачи

Современный этап развития теории алгоритмов в части разработки методов и алгоритмов поиска по образцу связан с детализацией описания эталонных конструкций, по которым выполняется поиск.

В исследовании в отличие от известных формализаций вводятся сущности – простой и составной образец. Такая детализация позволяет более гибко описывать взаимосвязи между элементами и задавать различные отношения на них (предпочтение, следование, альтернатива и др.).

Простой образец – дискретный объект в виде последовательности букв алфавита, размер простого образца определяется количеством букв, в него входящих. При этом логическое и физическое расстояния между смежными буквами совпадают и определяются метрикой задания образца. Такие типовые структуры данных как строка, список, множество (куча) букв описывают простые образцы. Примеры простых образцов приведены на рис. 1. В алфавите $A = \{a, b, c, d\}$ образец задан в виде строки (а), списка (б) и множества (с).

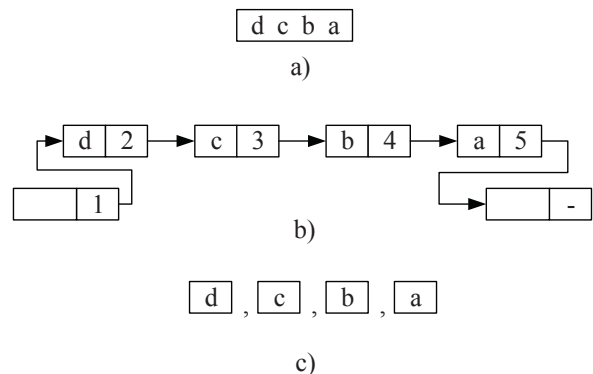


Рис. 1. Примеры представления простого образца

Составной образец – дискретный объект в виде последовательности букв алфавита, размер которого определяется количеством букв, в нее входящих, и метасимволов, указывающих на различные отношения между символами. При этом логическое и физическое расстояния между смежными буквами не совпадают. Структура составного образца характеризуется распределенными позициями смежных символов по его длине.

Для описания составных образцов используются сложные конструкции данных, имеющие сетевое представление. Элементами сети выступают преимущественно типовые структуры данных. Пример составного образца приведен на рис. 2. В алфавите $A = \{a, b, c, d\}$ составной образец задан в виде комбинации списка и подстрок различной длины, где \$ – подстрока переменной длины.

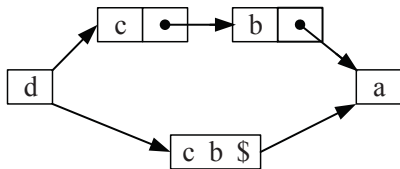


Рис. 2. Пример представления составного образца

В общем случае рассматривается задача поиска вхождения составного образца в анализируемую последовательность (текст, набор элементов). Составной образец и текст представляют одномерные структуры данных, составленные из букв рабочего алфавита.

Пусть заданы составной образец x длиной m -символов, содержащий z подстрок $x_1, x_2 \dots x_z$ с длинами $t_1, t_2, \dots t_z$ такими, что $t_1 + t_2 + \dots + t_z \leq m$, и текст y длиной n символов, где $n > 0, m > 0$ и $m \leq n$. Требуется найти позиции вхождения подстрок из x в y , т.е. определить такие наименьшие адреса $i_1, i_2 \dots i_z$, при которых справедливо

$$\forall j | y(i, i + m - 1) = x(1, t_j), j = 1 \dots z, (i_1 < i_2 < \dots i_z). \quad (1)$$

В частном случае, для образцов, включающих только одну подстроку, задача поиска формулируется как поиск вхождения простого образца. При этом вторая часть условия (1) записывается как

$$i_1 + 1 = i_2; i_2 + 1 = i_3; \dots i_{z-1} + 1 = i_z. \quad (2)$$

По умолчанию, выражение (2) при поиске по образцу подразумевается. Оно определяет физически смежный характер расположения символов из x в структуре текста y .

При $z=1$ задача поиска вхождения образца x длиной m символов в текст y длиной n символов имеет вид [12]

$$\forall i | (x(1, m) = y(i, i + m - 1)) | 1 \leq i \leq k, k = n - m + 1. \quad (3)$$

Другим практически значимым частным случаем является редукция составного образца x , к образцу, содержащему z подстрок $x_1, x_2 \dots x_z$ с длинами, равными одному символу каждая, т.е. $t_1 = t_2 = \dots = t_z = 1$ [13]. Примеры составных образцов приведены на рис.3. Здесь в расширенном алфавите $A = \{a, b, c, d, *\}$ составной образец задан как последовательность рабочих символов $\{a, b, c, d\}$ и метасимвола «*», указывающего переменную часть составного образца (начало или тело или окончание).

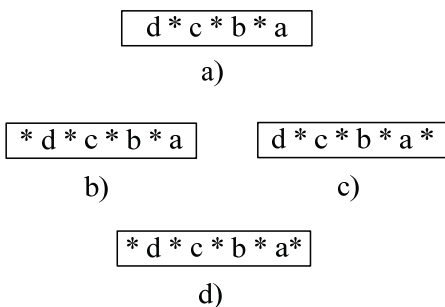


Рис. 3. Примеры редуцированного составного образца

Данная форма представления составного образца (рис.3) отличается тем, что расстояние между соседними символами образца x не всегда равно 1, что задает разнообразие в описании образца и позволяет более гибко вести поиск вхождений. Фактически в тексте y осуществляется посимвольный поиск из x , для каждого символа которого не обязательно выполняется условие (2), но обязательно выполняется условие (1).

Метод решения

Известные методы и алгоритмы последовательного поиска по образцу основаны на направленном последовательном сравнении текущих символов двух символьных операндов. Использование подготовленной таблицы переходов по результатам пересечений префиксов образца позволяет исключить полный перебор. Значения индексов возврата в таблице переходов, полученные на этапе предобработки образца, обеспечивают уменьшенный отступ по длине анализируемого текста при обнаружении текущего отвергающего символа образца. Алгоритмы Бойера-Мура, Кнута-Морриса-Пратта, Хорспула и др. являются наиболее известными алгоритмами, использующие идею предобработки [9,14].

Тем не менее, временные затраты последовательных методов и алгоритмов, в общем случае, имеют квадратичную сложность, что не допустимо для задач большой размерности с динамично меняющейся обстановкой. Как правило, данные методы и алгоритмы на их основе имеют программную реализацию, так как используют циклические конструкции для направленного перебора частичных вхождений префиксов образца и текущих символов в анализируемый текст с использованием дополнительной информации в виде таблицы переходов.

Известные аппаратно-ориентированные методы и основанные на них алгоритмы развивают идеи конвейерного и матричного поиска.

Алгоритмы конвейерного сопоставления двух символьных операндов (слов) основаны на вычислении, хранении и последовательной обработке двух двоичных (характеристических) векторов на каждом шаге сопоставления [15, 16, 17]. Каждый бит текущего характеристического вектора (ХВ) принимает значение логической «1» или «0» в зависимости от того, присутствует ли текущий символ первого слова-операнда в позиции второго слова-операнда. Алгоритмизация процессов сопоставления сводится к циклической обработке предыдущего и текущего ХВ. Элементарная операция обработки двух векторов представляет собой побитовое логическое умножение двоичных векторов со смещением предыдущей и текущей позиций. Пример конвейерного поиска образца O в тексте T с построчной обработкой ХВ приведен на рис. 4 [18].

Алгоритмы конвейерного сопоставления символьных операндов (строк) основаны на получении и последовательной обработке характеристических (двоичных) векторов, характеризующих в рамках отдельного вектора позиционные односимвольные вхождения текущего символа во второй операнд (текст или образец).

Недостаток данных конвейерных алгоритмов заключается в последовательной обработке двоичных векторов, что приводит к непродуктивным затратам времени для итерационного вычисления конечного характеристического вектора в соответствии с временными отсчетами сопоставлений.

образец <i>O</i>	XB	текст <i>T</i> A B A B A S K	Моменты времени сравнений
A B A	XB ₀	1 1 1 1 1 1 1 1	1
	XB ₁	0 1 0 1 0 1 0 0	
	XB ₂	0 0 1 0 1 0 0 0	3
	XB ₃	0 0 0 1 0 1 0 0	

Рис. 4. Конвейнерный поиск вхождений с построчной обработкой XB

Известен метод матричного поиска простого образца [1,19,20], основанный на двумерном представлении сопоставлений символов двух строк и заключающийся в организации вычислений по диагоналям двоичной (характеристической) матрицы. Характеристическая матрица поисковых ячеек состоит из $n \times k$ поисковых ячеек ($k=m-n+1$ – количество диагоналей в матрице). Она имеет геометрическую форму параллелограмма, в каждой строке которого располагается k поисковых ячеек, сдвинутых относительно следующей строки ячеек вправо на 1 позицию, начиная с ячеек первой строки. Такая форма матрицы использует принципы коллективного доступа к элементам, распределенным по диагоналям матрицы, и параллельной обработки диагональных элементов, свойственные ассоциативной памяти. Поиск образцов проходит по элементам k диагоналей, начиная от первой строки матрицы к ее последней строке с раздельным вычислением стартовых значений ячеек.

Алгоритм параллельного поиска простого образца показан на рис. 5. Блоки A1, A2, A4 являются обеспечивающими вычисления. Поисковые вычисления реализуются в блоке параллельной обработки k двоичных замечаний на основе $A3_{11} - A3_{nm}$, имеющих соединения в виде параллелограмма.

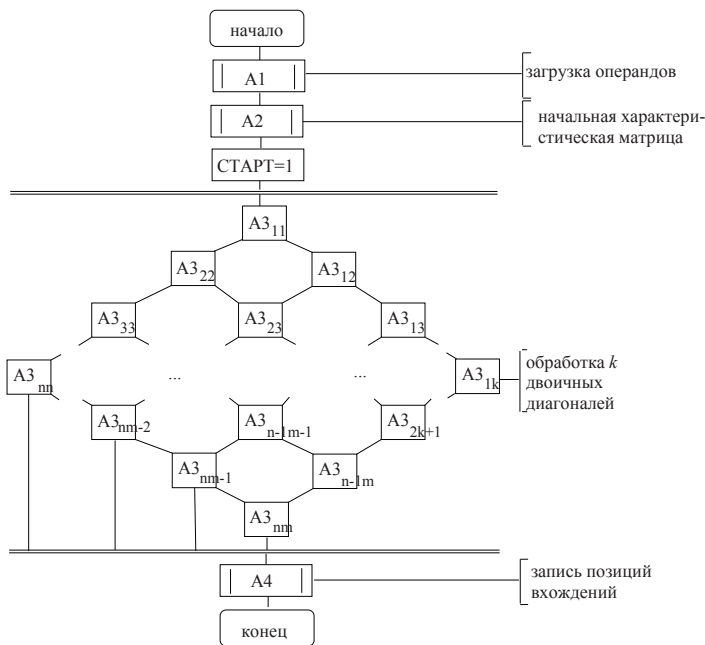


Рис. 5. Общий вид алгоритма матричного поиска простого образца

Параллельная обработка k диагоналей матрицы позволяет снять ограничения конвейерного поиска. Тем не менее, при организации поиска составного образца данный метод имеет функциональный недостаток. В нем невозможно учесть нерегулярно распределённые позиции отдельных символов составного образца в тексте, поэтому требуется дополнительная проверка текущей и предыдущей позиций односимвольных вхождений в характеристической матрице по условию из (1). Как следствие, появляются избыточные затраты времени на последовательную обработку каждого символа составного образца.

Известен способ поиска составного образца в последовательности [13], отличающийся параллельной обработкой элементов в составе k диагоналей характеристической матрицы и построчным вычислением стартовых значений ячеек матрицы, что позволяет снять необходимость проверки условия из (1). В каждой строке характеристической матрицы содержится цепочка k двухвходовых элементов ИЛИ, вычисляющих по строке стартовое значение для поиска составного образца. Тем не менее, данный способ не позволяет вести одновременный поиск простого и составного образцов вследствие специализированного характера связей ячеек характеристической матрицы, поддерживающей только нерегулярно распределённые позиции следования символов составного образца в тексте.

Пример матричного поиска составного образца показан на рис. 6 для $x=a*b*c$ и $y=caccbbcbcabcs$ для алфавита $A=\{a, b, c, d\}$.

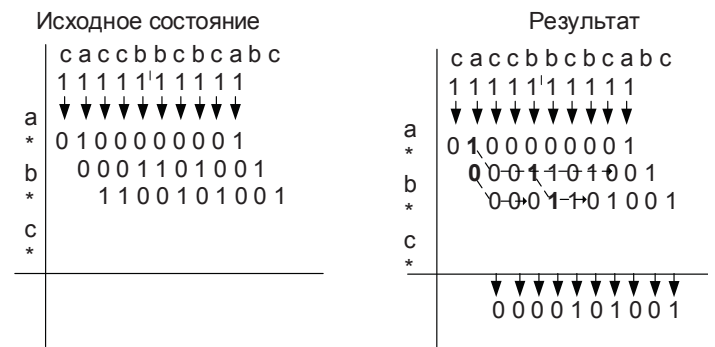


Рис. 6. Пример матричного поиска составного образца

Предлагаемый метод матричного поиска простого и составного образцов состоит из следующих шагов вычисления:

- шаг задания стартовых значений поиска простого образца
- шаг построчного вычисления стартовых значений составного образца;
- шаги параллельной обработки ячеек k диагоналей характеристической матрицы с учетом диагонально регулярных и построчно нерегулярных связей ячеек;
- шаг фиксации результатов поиска в виде двух двоичных векторов, биты которых показывают вхождения простого и составного образцов.

Предлагаемый метод матричного поиска простого и составного образцов отличается, во-первых, введением связей между поисковыми ячейками в составе k диагоналей и, во-вторых, заданием начального k -битового характеристического вектора со значением 11...11.

Этот вектор подается на поисковые ячейки первой строки характеристической матрицы для организации параллельного поиска простого образца. Метод матричного поиска простого и составного образцов является аппаратно-ориентированным. Для поддержки всех базовых вычислительных операций разработано специализированное однородное устройство, операционная часть которого содержит (рис. 7):

- характеристическую матрицу $n \times k$ поисковых ячеек;
- регистры $Pgx_1 - Pgx_n$ для хранения символов образца;
- регистры $Pgy_1 - Pgy_m$ для хранения символов текста;
- два массива триггеров $TP1_1 - TP1_k$, $TP2_1 - TP2_k$, для хранения позиций вхождений простого и составного образцов.

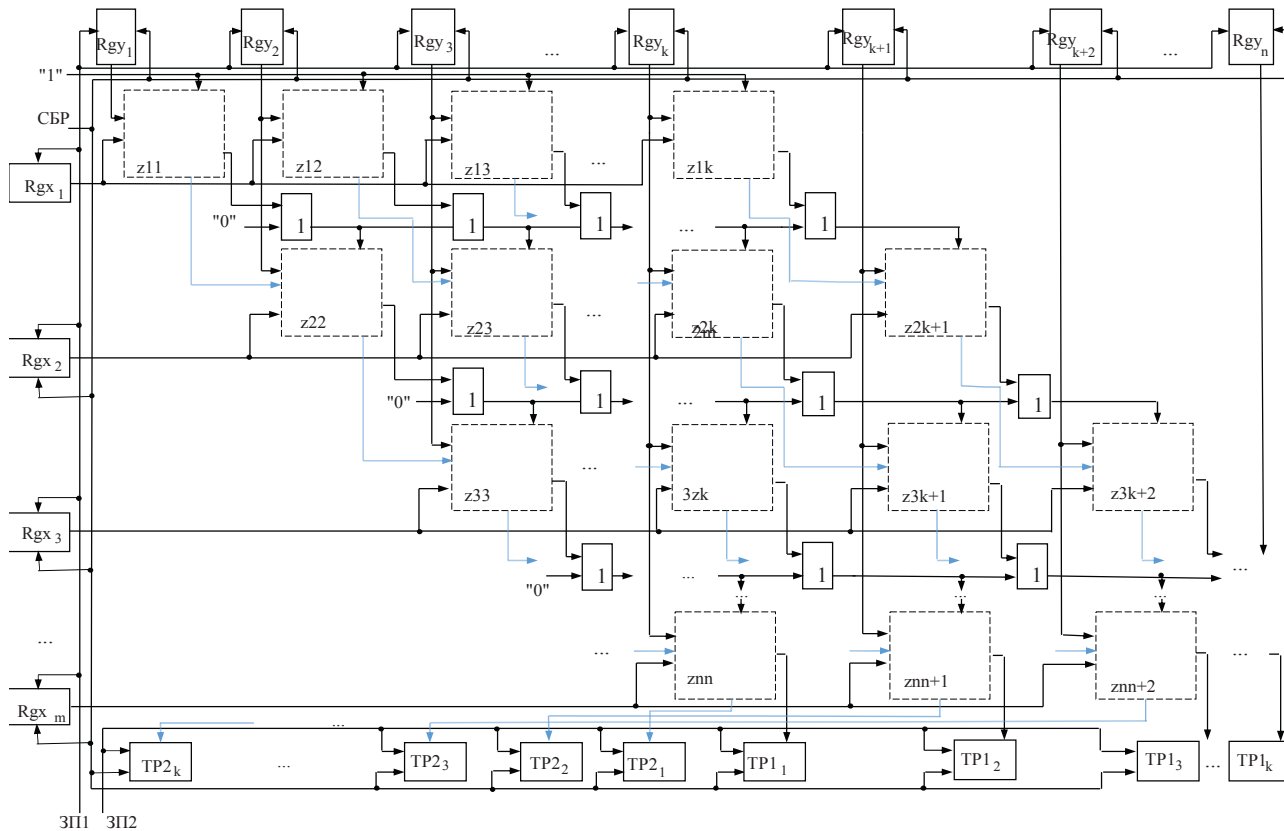


Рис. 7. Характеристическая матрица поиска простого и составного образцов

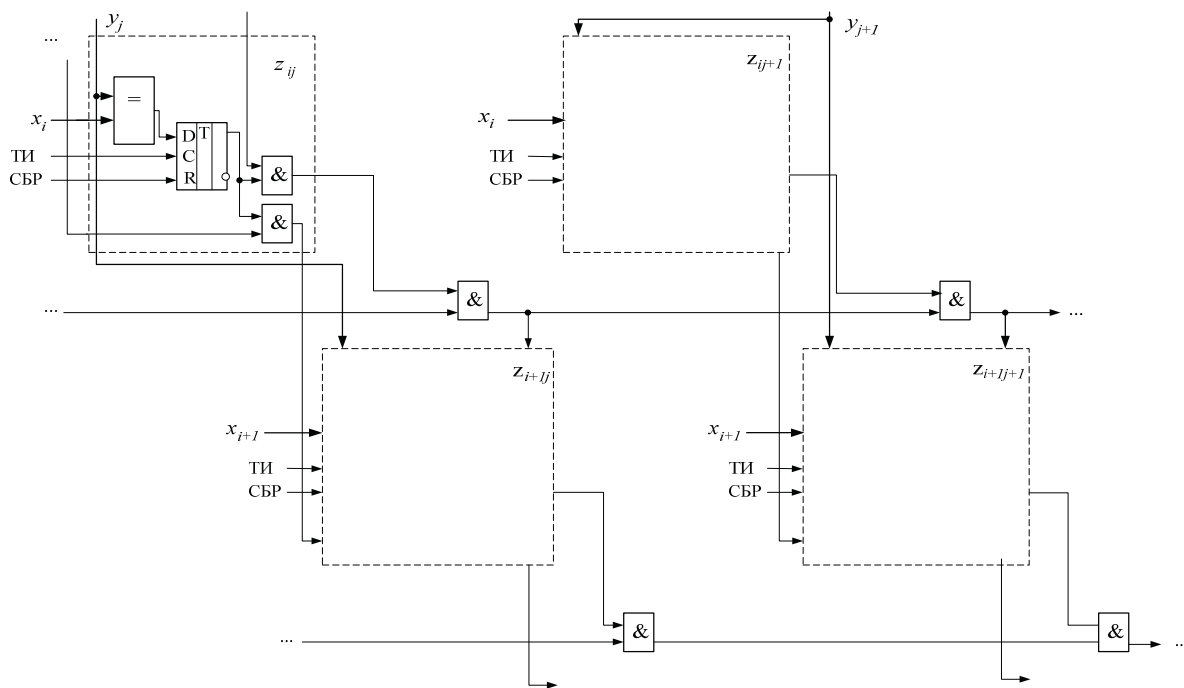


Рис. 8. Фрагмент характеристической матрицы

Кроме входов подачи p -разрядных кодов символов образца x_i и текста y_j , управляющих входов сброса СБР, тактирования ТИ, каждая ячейка z_{ij} матрицы содержит два входа поиска простого и составного образцов. Первый вход поиска непосредственно соединяет смежные ячейки в пределах диагоналей для поиска простого образца. Начальные значения для поиска простого образца задаются в первую строку матрицы k -разрядным вектором $11\dots 11$. Второй вход поиска необходим для построчного вычисления начальных значений для поиска составного образца. Эти значения вычисляются с помощью последовательности двухвходовых элементов И, фиксирующих в текущей строке положительный результат сравнения пары символов (образца и текста). Вычисляемые с помощью последовательности двухвходовых элементов И значения подаются в соответствующие позиции следующей строки, задавая тем самым поиск следующего символа составного образца.

Фрагмент матрицы однородного устройства поиска простого и составного образцов, содержащий 4 ячейки $z_{ij} - z_{i+1j+1}$, приведен на рис. 8, где ТИ – тактовый импульс, СБР – сигнал сброса.

Обсуждение и результаты

Согласно рис. 8 выходные поисковые функции ячейки имеют максимальную задержку, равную $T = \tau_{\Sigma} + \tau_{\Gamma}$. триггер $+ \tau_{2И} = 5\tau_{2И}$. Данная задержка определяет предельную частоту работы характеристической матрицы [21-22]. При этом параллельные процессы обработки смежно-диагональных ячеек в k диагоналях матрицы конкурируют с последовательными процессами построчного вычисления стартовых значений для ячеек, распределенных по матрице с учетом отношения следования символов. Первые процессы необходимы для поиска простого образца, вторые – для поиска составного образца.

Для оценки работоспособности метода и эффективности однородного устройства одновременного поиска простого и составного образцов было выполнено моделирование работы следующих матричных методов:

- матричный метод поиска простого образца (метод ППО);
- матричный метод поиска составного образца (метод ПСО);
- матричный метод поиска простого и составного образцов (метод ППО и ПСО).

Моделирование производилось при фиксированной длине текста $n=16$ символов и переменной длине образца $m=1\dots 15$ символов. Также моделировалась переменная длина текста $n=4\dots 24$ и фиксированная длина образца $m=2$. Данные сочетания позволили выявить предпочтительные области применения разработанного метода

Результаты сравнения трех методов приведены на рис. 9 и 10 соответственно (в усл. тактах).

Анализ результатов моделирования (рис. 9) показал, что разработанный метод не выгоден для коротких образцов (до 4-х символов), так как значительную долю времени поиска занимает построчное вычисление стартовых значений для поиска составного образца.

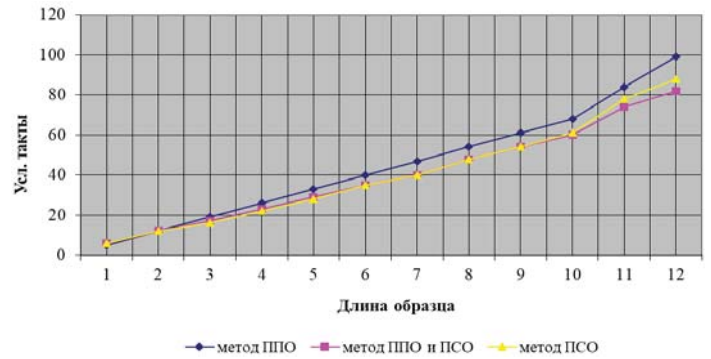


Рис. 9. Моделирование работы матричных методов от длины образца

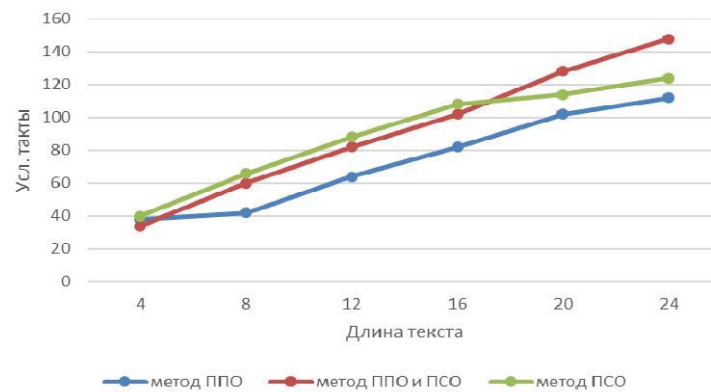


Рис. 10. Моделирование работы матричных методов от длины текста

Длина m образца более 4 символов увеличивает время обработки элементов на диагоналях характеристической матрицы, что нивелирует построчные вычисления и приводит к сбалансированному общему времени поиска двух типов образцов. В тоже время ранее известные методы матричного поиска становятся чувствительными к увеличению длины образца по затратам времени. Они проигрывают разработанному методу из-за избыточных затрат времени на многократную обработку символов образца по различным трассам в характеристической матрице.

Анализ результатов моделирования (рис. 10) показал, что разработанный метод выгодно использовать на текстах, длиной не более 16 символов, так как при увеличении этой длины значительная доля времени отводится на построчное вычисление стартовых значений для поиска составного образца, а затраты времени на обработку элементов диагоналей уменьшаются.

Увеличение целесообразной длины текста более 16 символов возможно с увеличением длительности периода тактового импульса, связанного со схемотехнической реализацией двух выходных поисковых функций и средней длительностью вычислений в ячейке матрицы, равной 5τ (усл. ед.). Тем не менее, данный путь ведет к снижению частоты работы специализированного однородного устройства, поэтому он не рассматривается для высокопроизводительных вычислительных систем.

Выводы

1. Для расширения функциональных возможностей поисковых методов введен термин «составной образец», позволяющий более гибко задавать отношения между соседними элементами в предметной области поиска.

2. Как развитие матричных методов поиска создан метод матричного поиска простого и составного образцов. Метод отличается одновременным учетом пространственного и временного расположения элементов образца в тексте, что позволяет задавать сложные информационно-логические отношения. Метод позволяет вести одновременный поиск простого и составного образцов за счет вычисления по элементам диагоналей двух выходных поисковых функций на основе раздельного вычисления для них стартовых значений.

3. На основе метода разработано однородное устройство поиска простого и составного образцов, определены рациональные характеристики размеры данных для его эффективной работы. Показано, что используя длительность вычислений выходных функций в рамках одной ячейки, определены предпочтительные длины образца (более 4-х символов) и текста (не более 16 символов). Устройство имеет линейный алгоритм управления, так как все базовые вычислительные операции имеют аппаратную поддержку.

4. Разработанные матричный метод и однородное устройство могут использоваться в процессорах логического вывода, интеллектуальных анализаторах баз данных и знаний, в системах естественно-языковой обработки информации и иных высокопроизводительных системах обработки данных с оперативным откликом на динамично меняющуюся обстановку [23-25].

Литература

1. *Огнев И.В., Борисов В.В., Сутула Н.А.* Ассоциативные память, среды, системы. М.: Горячая линия – Телеком, 2016. 420 с.
2. *Огнев И.В., Борисов В.В.* Интеллектуальные системы ассоциативной памяти. М.: Радио и связь, 1996. 176 с.
3. *Зерин И.С., Атакищев О.И., Титенко Е.А., Корольков О.Ф., Атакищев А.О.* Метод, алгоритм и техническое решение параллельного поиска и подстановки на ассоциативной памяти // В мире научных открытий. 2012. № 1-1 (25). С. 166-180.
4. *Мирталибов Т.А., Фролов С.Н., Ханис А.Л., Титенко Е.А.* Многоагентные системы в технической диагностике сложных технических объектов // Известия Юго-Западного государственного университета. 2015. № 3 (60). С. 18-25.
5. *Курочкин А.Г., Емельянов С.Г., Титенко Е.А.* Анализ алгоритмов поиска по образцу для управления группой роботов // Научное издание. 2014. Т. 15. № 12. С. 4-8.
6. *Каляев А.В., Левин И.И.* Модульно-наращиваемые многопроцессорные системы со структурно-процедурной организацией вычислений. М.: Янус-К, 2003. 380 с.
7. *Люгер, Дж. Ф.* Искусственный интеллект: стратегии и методы решения сложных проблем. М.: Издательский дом «Вильямс». 2003. 864 с.

8. *Марков А.А.* Избранные труды. Т. 2. Теория алгоритмов и конструктивная математика, математическая логика, информатика и смежные вопросы. М.: МЦНМО. 2007. 650 с.

9. *Дж. Макконелл.* Основы современных алгоритмов. М.: Техносфера, 2006. 368 с.

10. *Wichert A.* Artificial intelligence and a universal quantum computer. // AI Communications. Vol. 29. Issue 4. 2016. P. 537-543.

11. *Graham A. Stephen.* String searching algorithms / A. Stephen Graham. – Lecture Notes Series on Computing. Vol. 6. Singapore: World Scientific Publishing, 1994.

12. *Титенко, Е.А.* Метод и однородное вычислительное устройство k-приближительного поиска вхождений по образцу // Вестник Воронежского государственного технического университета. Т. 7. №7. 2011 С. 70-78.

13. Патент 2549525 РФ, МПК G06 F 12/00, G06 F 17/30. Способ и устройство поиска составного образца в последовательности / А.В. Крипачев, Е.А. Титенко, Р.В. Бредихин, Белокопытов А.В., Курочкин А.Г.; заявитель и патентообладатель Юго-Зап. гос. ун-т.: заявл. 15.07.2013; опубл. 27.04.2015, Бюлл. №12.

14. *Максимов В.* Алгоритмы поиска, или как искать неизвестно что // Монитор. 1995. №6. С. 10-16.

15. *Кулик Б.А.* Системы поиска в произвольном тексте // Программирование. 1987. №1. С. 6-10.

16. *Wu S., Manber U.* Fast text searching allowing errors, Commun. ACM. 35(10):83-91.

17. *Titenko E.A., Degtyarev S.V.* Approximate search in the sample on the basis Manber-Wu method // Journal of Fundamental and Applied Sciences. 2017. Vol. 9. No. 2. P. 914.

18. *Титенко Е.А.* Метод параллельного поиска по образцу и матричное устройство для его реализации // Информационные системы и технологии. №4. 2011. С. 24-30.

19. Патент 72771 Российская Федерация, МПК G06 F 12/00. Устройство для параллельного поиска и обработки данных / Е.А. Титенко, Л.А. Лисицин, В.М. Довгаль; заявитель и патентообладатель Курск. гос. техн. ун-т. - №2007149075/22; заявл. 25.12.2007; опубл. 27.04.2008, Бюлл. №12.

20. *Титенко Е.А.* Аппаратно-ориентированный алгоритм и однородное вычислительное устройство сопоставления слов // Вестник компьютерных и информационных технологий. №11. 2011. С. 51-56.

21. *Потемкин И.С.* Функциональные узлы цифровой автоматизации. М.: Энергоатомиздат. 1988. 320 с.

22. *Новожилов О.П.* Основы цифровой техники. Учебное пособие. М.: ИП Радиософт, 2013. 528 с.

23. *Мельник Э.В., Клименко А.Б., Иванов Д.Я., Гандурин В.А.* Методы обеспечения бесперебойной работы сетевых информационных-вычислительных систем с кластеризацией // Известия Южного федерального университета. Технические науки. 2016. № 12 (185). С. 71-84.

24. *Винограденко А.М., Будко Н.П.* Адаптивный контроль технического состояния автономных сложных технических объектов на основе интеллектуальных технологий // Т Comm: телекоммуникации и транспорт. Том. 14. № 1. 2020. С. 25-35.

25. *Довбня В.Г., Фролов С.Н., Сулима К.П., Щитов А.Н.* Специфика Реализаций комплексов управления на базе технологии LORAWA // Т Comm: Телекоммуникации и транспорт. Том. 14. № 9. 2020. С. 24-30.

THE METHOD AND MULTIPLE DEVICE OF A MATRIX SEARCH FOR SIMPLE AND COMPLEX SAMPLES

Evgeny A. Titenko, The Southwest State University, Kursk, Russia, johnnit@mail.ru
Alexey N. Schitov, The Southwest State University, Kursk, Russia, a.n.schitov@mail.ru

Abstract

The research consists in the development of hardware pattern search methods that use the principles of shared access and parallel data processing inherent in associative memory. Reducing the search time is achieved through the formation and parallel processing of a binary (characteristic) matrix of comparisons of pattern symbols and text. A composite pattern has been introduced, it allows flexible description of search terms. The characteristic matrix is in the form of a parallelogram; it consists of rows shifted to the right, starting from the first row. This form allows you to conduct a parallel search for simple and composite patterns on the diagonal elements of the matrix. The developed method supports hardware search in the characteristic matrix. It is distinguished by the simultaneous consideration of local and distributed relationships between the elements of the diagonals and rows of the matrix. Separate calculation of starting values along the diagonals of the matrix allows you to independently search for two types of samples by calculating in the cells of the diagonals of the characteristic matrix of two output search functions. The method has linear time and quadratic hardware complexity. The clock pulse duration is determined by the sum of the delays of the comparison circuit for a pair of symbols, a D-flip-flop and a two-input element I. The homogeneous structure of the matrix search device, the use of standard operations allow the device to be implemented on a promising FPGA element base, which determines its use in high-performance systems for processing and transmitting heterogeneous information.

Keywords: characteristic matrix, diagonal search, occurrence, string operand

References

1. I.V. Ognev, V.V. Borisov, N.A. Stool (2016). Associative memory, environments, systems. Moscow: Hot line – Telecom. 420 p.
2. I.V. Ognev, V.V. Borisov (1996). Intelligent systems of associative memory. Moscow: Radio and communication. 176 p.
3. I.S. Zerin, O.I. Atakischev, E.A.Titenko, O.F. Korolkov, A.O. Atakischev. (2012). Method, algorithm and technical solution of parallel search and substitution on associative memory. *In the world of scientific discoveries*. No. 1-1 (25). P. 166-180.
4. T.A. Mirtalibov, S.N. Frolov A.L. Khanis, E.A. (2015). Titenko Multiagent systems in technical diagnostics of complex technical objects. *News of the South-West State University*. No. 3 (60). P. 18-25.
5. A.G. Kurochkin, S.G. Emelyanov, E.A. Titenko (2014). Analysis of pattern-based search algorithms for controlling a group of robots. *Science-intensive technologies*. Vol. 15. No. 12. P. 4-8.
6. A.V. Kalyaev, I.I. Levin (2003). Modularly expandable multiprocessor systems with structural and procedural organization of computations. Moscow: Yanus-K. 380 p.
7. J.F. Luger (2003). Artificial intelligence: strategies and methods for solving complex problems. Moscow: Williams Publishing House. 864 p.
8. A.A. Markov (2007). Selected Works. vol. 2. Theory of algorithms and constructive mathematics, mathematical logic, computer science and related issues. Moscow: MTsNMO. 650 p.
9. J. McConnell (2006). Fundamentals of modern algorithms. Moscow: Tekhnosfera. 368 p.
10. A. Wichert (2016). Artificial intelligence and a universal quantum computer. *AI Communications*. Vol. 29. Issue 4. P. 537-543.
11. A. Stephen Graham (1994). String searching algorithms. *Lecture Notes Series on Computing*. Vol. 6. Singapore: World Scientific Publishing.
12. E.A. Titenko (2011). Method and homogeneous computing device of k-approximate search for occurrences by pattern. *Bulletin of the Voronezh State Technical University*. Vol. 7. No. 7. P. 70-78.
13. Patent 2549525 RF, IPC G06 F 12/00, G06 F 17/30. Method and device for searching a composite sample in a sequence / A.V. Kripachev, E.A. Titenko, R.V. Bredikhin, Belokopytov A.V., Kurochkin A.G. ; applicant and patentee Yugo-Zap. state un-t.: Appl. 07/15/2013; publ. 04/27/2015, Bull. No. 12.
14. V. Maksimov (1995). Search algorithms, or how to search for unknown what *Monitor*. No. 6. P. 10-16.
15. B.A. Kulik (1987). Search systems in free text. *Programming*. No. 1. P. 6-10.
16. S. Wu, U. Manber. Fast text searching allowing errors, *Commun. ACM*. 35 (10): 83-91.
17. E.A. Titenko, S.V. Degtyarev (2017). Search in the sample on the basis Manber-Wu method. *Journal of Fundamental and Applied Sciences*. Vol. 9. No. 2. P. 914.
18. E.A. Titenko (2011). Parallel pattern search method and matrix device for its implementation. *Information systems and technologies*. No. 4. P. 24-30.
19. Patent 72771 Russian Federation, IPC G06 F 12/00. Device for parallel search and data processing / E.A. Titenko, L.A. Lisitsin, V.M. Dovgal; applicant and patentee Kursk. state tech. un-t. No. 2007149075/22: app. 12/25/2007; publ. 27.04.2008, Bull. No. 12.
20. E.A. Titenko (2011). Hardware-oriented algorithm and homogeneous computing device for comparing words. *Bulletin of computer and information technologies*. No. 11. P. 51-56.
21. I.S. Potemkin (1988). Functional units of digital automation. Moscow: Energoatomizdat. 320 p.
22. O.P. Novozhilov (2013). Fundamentals of digital technology. Tutorial. Moscow: IP Radiosoft. 528 p.
23. E.V. Melnik, A.B. Klimenko, D.Ya. Ivanov, V.A. Gandurin (2016). Methods for ensuring uninterrupted operation of network-centric information-computing systems with clustering. *News of the Southern Federal University. Technical science*. No. 12 (185). P. 71-84.
24. A.M. Vinogradenko, N.P. Budko (2020). Adaptive control of the technical state of autonomous complex technical objects based on intelligent technologies. *T-Comm*. Vol. 14. No. 1. P. 25-35.
25. V.G. Dovbnya, S.N. Frolov, K.P. Sulima, A.N. Shchitov (2020). Specificity of Implementation of control systems based on LORAWA technology. *T-Comm*. Vol. 14. No. 9. P. 24-30.