

ПРИМЕНЕНИЕ ССЫЛОЧНОГО РАНЖИРОВАНИЯ ПРИ МОДЕЛИРОВАНИИ ПАРАЛЛЕЛЬНЫХ СПЕЦИАЛИЗИРОВАННЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ НА ТРАНСПОРТЕ



А.С. Филипченко

Описываются математические принципы работы алгоритма ссылочного ранжирования на основе синтеза линейной алгебры и теории графов. Выполняется адаптация алгоритма к графовой модели топологии параллельных специализированных вычислительных систем. Для апробации предложенного метода моделирования весовых характеристик узлов вычислительных систем используется ориентированный граф параллельной специализированной вычислительной системы «CDF 9/7». Демонстрируется программная реализация адаптированного алгоритма ссылочного ранжирования в виде веб-приложения с графическим интерфейсом.

Ключевые слова: моделирование, ранжирование, распараллеливание, топология, графы, алгоритм, процессоры

EDN: BMHCSJ

При моделировании параллельных специализированных вычислительных систем на транспорте актуальной задачей является определение весомости процессорных элементов. В рамках данной задачи требуется определить, какие из процессорных элементов многопроцессорной вычислительной системы являются более «важными», а какие – менее «важными». Ответ на этот вопрос позволит организовать эффективную неодноранговую систему приоритизации распределения вычислительных ресурсов внутри вычислительной системы.

При этом топология связей между процессорными элементами в вычислительной системе может быть представлена в виде графа. В этой интерпретации связи между процессорными элементами могут быть представлены как ребра графа, а сами процессорные элементы – как вершины [1;2].

В то же время существует алгоритм ссылочного ранжирования страниц, который используется,

например, в поисковых системах сети Интернет для подготовки списка релевантной выдачи на запрос. Такой алгоритм называется «PageRank» и используется, например, в поисковой системе Google [3]. Алгоритм применяет методы линейной алгебры к графу веб-страниц Интернета и возвращает список их приоритетов.

В рамках настоящей работы ставится цель адаптировать алгоритм ранжирования страниц «PageRank» к использованию в сфере моделирования параллельных специализированных вычислительных систем.

Принципы работы алгоритма ссылочного ранжирования

Алгоритм ссылочного ранжирования применяется на практике при решении задач, актуальных для транспортной отрасли. Например, алгоритм применяется в задаче выделения ролей остановок сетей общественного транспорта в Санкт-Петербурге [4].

Филипченко Александр Сергеевич, аспирант кафедры «Вычислительные системы, сети и информационная безопасность» Российского университета транспорта (РУТ (МИИТ)), лауреат именной стипендии имени Заслуженного деятеля науки и техники РСФСР, д.т.н., профессора Кочнева Ф.П. за высокие достижения в учебе. Область научных интересов: методы моделирования параллельных специализированных вычислительных систем на транспорте. Автор 22 научных работ. Имеет четыре патента на изобретения.

Взвешенные связи интегрируют информацию о расстоянии и количестве пересадок в маршрутах движения между станциями [4]. Алгоритмы ссылочного ранжирования также используются в интеллектуальной поддержке принятия решений при формировании программ тренажерной подготовки и производственных полетов и на основе методов семантического анализа текстов на основе нейросетевых моделей языка [5]. Еще алгоритм «PageRank» используется на транспорте в качестве вероятностного метода решения задачи построения матриц корреспонденций [6]. Метод находит свое применение в исследованиях, связанных с количественными, структурными и параметрическими изменениями транспортных потоков в городской транспортной системе [6].

Алгоритм «PageRank» дает точную оценку веса страницы и помогает ранжировать страницы для выдачи их в качестве результата в ответ на поисковый запрос [7]. В основе алгоритма лежит представление, что если страница А ссылается на страницу В, значит, страница А считает, что страница В — важная страница [8]. Ранги веб-страниц вычисляются в рамках цикла. Алгоритм предполагает возврат и пересчет всех рангов в рамках каждой итерации, так как «PageRank» остальных веб-страниц, ссылающиеся на веб-страницу А, уже изменится. В процессе разработки алгоритма «PageRank» создателям Google пришлось доказать эргодическую теорему, которая заключается в том, что процесс пересчета рангов веб-страниц в «конечном» итоге сойдется [9].

Алгоритм «PageRank» моделирует процесс случайного перемещения пользователя по страницам веб-сети. При этом перед каждым новым шагом с заданной вероятностью выбирается одно из двух действий [10]:

- пользователь с равной вероятностью выбирает одну из гиперссылок, представленных в виде ребер графа и принадлежащих текущей странице, и переходит на страницу, являющуюся вершиной графа, по выбранной ссылке;
- пользователь осуществляет перемещение на другую страницу, выбранную случайным образом из всех существующих страниц веб-сети.

Но каким образом передать граф на вход алгоритма, со всеми его вершинами, ребрами и их направлениями? Для решения этой задачи существует метод алгебраической формализации графовых структур данных путем преобразования их в матрицы смежности. Матрица смежности графа G с конечным числом вершин n определяется как квадратная матрица M размерностью $n \times n$, в которой значение элемента a_{ij} обозначает вес ребра из i -й вершины в j -ю. А в качестве веса ребра ij имеет смысл взять вероятность

перехода P_{ij} из вершины i в вершину j , которая представляется следующим тождеством (1).

$$P_{ij} = \frac{1}{n_i} = n_i^{-1}, \quad (1)$$

где n_i — число ребер, исходящих из вершины i .

Для решения проблемы вершин-петель в алгоритме «PageRank» предусмотрена функция перемещения [10], которая определяет, что в каждой итерации метка может:

- выполнить переход по случайному ребру с вероятностью β ;
- выполнить переход к случайной вершине с вероятностью $(1-\beta)$.

Величина β обычно выбирается в интервале [0,8; 0,9]. Для реализации функции перемещения необходимо подготовить дополнительную матрицу смежности TP , имитирующую полносвязную топологию исходного графа. Такая матрица отвечает за выполнение перехода к случайной вершине. Чтобы получить модифицированную матрицу смежности M' с поддержкой функции перемещения необходимо выполнить преобразование (2).

$$M' = M \times \beta + TP \times (1-\beta). \quad (2)$$

Далее в рамках алгоритма необходимо определить вектор начального распределения вероятностей, каждый элемент которого будет представлять вероятность перехода в соответствующую вершину графа. Обычно данный вектор инициализируется равномерным распределением вероятностей между всеми вершинами графа. Если матрицу переходных вероятностей представить как марковскую цепь, то динамика такой цепи с тактовым временем t будет описываться векторно-матричным рекуррентным уравнением (3).

$$P(t+1) = M' \times P(t), \quad (3)$$

где $P(t)$ обозначает вектор вероятностей переходов в момент времени t , а $P(t+1)$ — вектор вероятностей переходов на следующем шаге $t+1$ [11].

По результатам этого преобразования мы получим вектор той же размерности, но уже с другим распределением вероятности перехода к вершинам внутри него. Этот вектор будет отображать ранги вершин после прохождения t -й итерации по графу, представленному в виде матрицы смежности. Соответственно, чем дольше мы будем моделировать блуждание метки по графу, тем точнее мы получим вектор вероятностей. Таким образом, полученный вектор нам необходимо снова умножить на матрицу, чтобы получить новый, более точный вектор, чтобы снова умножить его на матрицу. Данный цикл требуется продолжать до тех пор, пока разница между соседними

векторами не достигнет бесконечно малых значений ϵ . В этот момент мы получим собственный вектор матрицы смежности графа, который и будет представлять устойчивое распределение вероятностей между вершинами. Именно этот собственный вектор и будет являться результатом работы алгоритма «PageRank».

Более того, вычисление собственных векторов в параметрической форме демонстрирует связь между их компонентами и древовидной структурой ориентированного графа [12]. Таким образом, чем выше вероятность перехода в вершину, тем выше ее ранг и тем она «важнее». От описанной теоретической базы перейдем непосредственно к задаче моделирования параллельных специализированных вычислительных систем на транспорте.

Графовая модель параллельной специализированной вычислительной системы на транспорте

Возьмем параллельную специализированную вычислительную систему, реализующую алгоритм «CDF 9/7». У алгоритма «CDF 9/7» имеется широкая сфера применения в транспортных задачах, в частности в обработке информации об окружающей обстановке [13], обнаружении препятствий на пути следования [14] и в сегментации железнодорожного пути [15]. Также алгоритм показывает хорошие показатели при исследовании его с использованием закона Амдала [2].

В процессе составления расписания параллельных вычислений улучшенной реализации алгоритма «CDF 9/7» путем модификации графа вычислений получен модифицированный граф связей между процессорными элементами в вычислительной системе [2]. Данный граф показан на рис. 1.

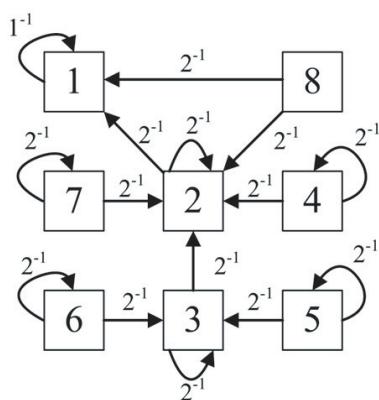


Рис. 1. Граф параллельной специализированной вычислительной системы «CDF 9/7» с взвешенными ребрами

Программная реализация ранжирования процессорных элементов в параллельной специализированной вычислительной системе

На основе описанного алгоритма ссылочного ранжирования на языке JavaScript разработано веб-приложение с графическим интерфейсом на HTML. При запуске программы отображает интерфейс, в котором пользователю необходимо ввести число процессорных элементов в исследуемой вычислительной системе. В зависимости от введенного числа программа формирует форму ввода матрицы смежности. Пример заполнения данной формы для графа нашей параллельной специализированной вычислительной системы приведен на рис. 2.

Нажатие на кнопку «Отправить» запустит процедуру нахождения собственного вектора матрицы смежности, после чего результаты ранжирования процессорных элементов будут выведены на экран в виде линейчатой диаграммы. Полученные в результате ранги процессорных элементов для параллельной специализированной вычислительной системы «CDF 9/7» показаны на рис. 3.

Данная программа для ЭВМ успешно прошла процедуру государственной регистрации в Роспатенте, получено соответствующее авторское свидетельство [16].

Заключение

В результате ссылочное ранжирование применено в задаче моделирования параллельных специализированных вычислительных систем на транспорте на примере вычислительной системы, реализующей алгоритм «CDF 9/7». Ранги процессорных элементов

Рис. 2. Заполнение матрицы смежности в программе ранжирования

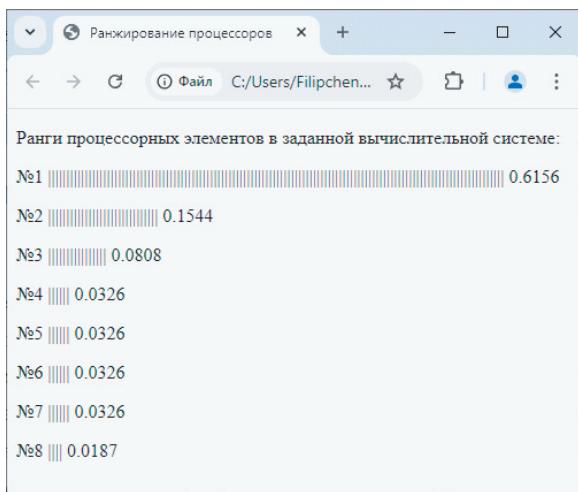


Рис. 3. Результат ранжирования процессорных элементов в параллельной специализированной вычислительной системе «CDF 9/7»

в данной модели разделились на три основных группы: высокие, средние и низкие.

Таким образом, можно сделать вывод, что в рамках обеспечения отказоустойчивости информационно-вычислительной инфраструктуры на транспорте, имеет смысл выполнить резервирование процессорных элементов путем дублирования в зависимости от их рангов. Данная задача может быть решена методом неопределенных множителей Лагранжа, в котором ранги будут представлены в виде стоимостей отдельных подсистем. Также сведения о рангах процессорных элементов позволят использовать неоднородную схему приоритизации распределения вычислительных ресурсов для увеличения производительности. Процессорным элементам с высокими рангами можно выделять больше ресурсов памяти и времени за счет процессорных элементов с низкими рангами.



Литература

1. Селиверстов, Е. Ю. Графовые модели графического процессора / Е. Ю. Селиверстов. - Текст : непосредственный // Системы компьютерной математики и их приложения. - 2017. - № 18. - С. 117-119.
2. Филипченко, А. С. Вычисление показателей эффективности оптимизированной реализации параллельного алгоритма CDF 9/7 / А. С. Филипченко. - Текст : непосредственный // Интеллектуальные транспортные системы : материалы III Международной научно-практической конференции, Москва, 30 мая 2024 г. - Москва : Российский университет транспорта (МИИТ), 2024. - С. 476-482. - DOI 10.30932/978-5002446094-2024-476-482.
3. Page, L. E. The PageRank Citation Ranking: Bringing Order to the Web / L. E. Page, S. M. Brin, R. Motwani, T. A. Winograd. - Текст : электронный // Proceedings of the 7th International World Wide Web Conference. - Brisbane, Australia. - 1998. - P. 161-172. - URL: <https://github.com/ermintham/Papers/blob/master/Google/Page%2CBrin%2CMotwani%2CWinograd-%20The%20PageRank%20Citation%20Ranking%20Bringing%20Order%20to%20the%20Web.pdf> (дата обращения: 19.11.2024).
4. Role discovery in node-attributed public transportation networks: the study of Saint Petersburg city open data / Yu. V. Lytkin, P. V. Chunaev, T. A. Gradov [et al.]. - DOI 10.17586/2226-1494-2023-23-3-553-563. - Текст : непосредственный // Scientific and Technical Journal of Information Technologies, Mechanics and Optics. - 2023. - Vol. 23, №. 3. - P. 553-563.
5. Перспективы применения технологий искусственного интеллекта при организации тренажерной подготовки пилотов в рамках концепции анализа фактических данных / А. Д. Худякова, Э. А. Болюта, Я. А. Жуткина [и др.]. - Текст : непосредственный // Теоретические и практические аспекты развития современной науки: теория, методология, практика : сборник научных статей по материалам VIII Международной научно-практической конференции, Уфа, 27 мая 2022 г. - Уфа: Общество с ограниченной ответственностью «Научно-издательский центр «Вестник науки», 2022. - С. 32-37.
6. Селиверстов, Я. А. Методы и модели построения матриц транспортных корреспонденций / Я. А. Селиверстов, С. А. Селиверстов. - Текст : непосредственный // Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Информатика. Телекоммуникации. Управление. - 2015. - № 2-3(217-222). - С. 49-70. - DOI 10.5862/JCSTCS.217-222.5.
7. Клюквин, Р. В. Способы определения ранжирования страниц. Алгоритм PageRank / Р. В. Клюквин, Ю. С. Белов. - Текст : непосредственный // В мире науки и инноваций : сборник статей международной научно-

практической конференции. В 5 частях, Казань, 20 апреля 2017 г. Ч. 4. - Казань: Общество с ограниченной ответственностью «Аэтерна», 2017. - С. 49-52.

8. Подлесный, А. О. Применение алгоритма PageRank для определения весов web-страниц / А. О. Подлесный, О. С. Полякова, С. А. Виденин. - Текст : непосредственный // Перспективное развитие науки, техники и технологий: сборник научных статей : материалы IV Международной научно-практической конференции, Курск, 17-18 октября 2014 г. / Председатель организационного комитета Горохов А. А. (ответственный редактор); Члены оргкомитета: Ивахненко А. Г., Сторублев М. Л. - Курск : Закрытое акционерное общество «Университетская книга», 2014. - С. 280-283.

9. Сарсенова, А. З. Алгоритм ссылочного ранжирования / А. З. Сарсенова. - Текст : непосредственный // Научное сообщество студентов: сборник материалов X Международной студенческой научно-практической конференции, Чебоксары, 17 июня 2016 г. - Чебоксары: Общество с ограниченной ответственностью «Центр научного сотрудничества «Интерактив плюс», 2016. - С. 136-138.

10. Фролов, А. С. Исследование подходов к реализации PageRank на языке параллельного программирования CHARM++ / А. С. Фролов, А. С. Семенов. - Текст : непосредственный // Современные информационные технологии и ИТ-образование. - 2016. - Т. 12, № 3-1. - С. 159-168.

11. Земсков, А. В. Аналитический подход к выборочному поиску функций вероятностей состояний в марковских цепях / А. В. Земсков. - Текст : непосредственный // Известия высших учебных заведений. Приборостроение. - 2023. - Т. 66, № 12. - С. 1035-1049. - DOI 10.17586/0021-3454-2023-66-12-1035-1049

12. Buslov, V. A. On the Characteristic Polynomial and Eigenvectors in Terms of the Tree-Like Structure of a Digraph / V. A. Buslov. - DOI: 10.1007/s10958-018-3854-5. - Текст : непосредственный // Journal of Mathematical Sciences. - 2018. - Vol. 232, № 1. - P. 6-20. - DOI: 10.1007/s10958-018-3854-5.

13. Савинов, К. Н. К вопросу инновационной железной дороги / К. Н. Савинов, М. Г. Яшин. - Текст : непосредственный // Инновационная железная дорога. Новейшие и перспективные системы обеспечения движения поездов. Проблемы и решения: сборник статей VI международной научно-практической конференции, Санкт-Петербург, Петергоф, 18 мая 2023 г. / Под общ. ред. М. Г. Яшина. - Санкт-Петербург, Петергоф: Военный институт (железнодорожных войск и военных сообщений) - структурное подразделение Федерального государственного казенного военного образовательного учреждения высшего образования «Военная академия материально-технического обеспечения им. генерала армии А.В. Хрулева» Министерства обороны Российской Федерации, 2023. - С. 6-15.

14. Попов, П. А. Поезд без машиниста - российские перспективы / П. А. Попов, А. Л. Охотников. - Текст : непосредственный // Автоматика, связь, информатика. - 2019. - № 8. - С. 4-6. - DOI 10.34649/AT.2019.8.001.

15. Мащенко, П. Е. Оптимизация модели нейронной сети U-HarDNet-70 для сегментации железнодорожного пути / П. Е. Мащенко, П. П. Ширяев. - Текст : непосредственный // Транспорт Российской Федерации. - 2020. - № 6(91). - С. 35-38.

16. Свидетельство о государственной регистрации программы для ЭВМ № 2024683722. Российская Федерация. Программа ранжирования процессорных элементов в параллельной специализированной вычислительной системе : № 2024681695 : заявлено 22.09.2024 : опубликовано 14.10.2024 / А. С. Филипченко. - Текст : непосредственный.