

УДК 004.021

ИССЛЕДОВАНИЕ ПАРАМЕТРОВ МУРАВЬИНОГО АЛГОРИТМА В КОНТЕКСТЕ ЗАДАЧИ ОПТИМИЗАЦИИ ВЫПОЛНЕНИЯ РАСПРЕДЕЛЁННЫХ ЗАПРОСОВ

Заславский В.А., Савкова Е.О.

*Донецкий национальный технический университет
кафедра автоматизированных систем управления
bobah22@gmail.com*

Рассмотрены параметры муравьиного алгоритма. Выбран способ исследования параметров и исходные данные для него. Проведено исследование, в результате чего были найдены наиболее оптимальные параметры алгоритма.

Актуальность задачи

4 SQL-ориентированные базы данных активно используются во многих современных приложениях. Работа с базами данных основана на использовании различных запросов, от скорости выполнения которых часто зависит скорость работы приложения. Проблема оптимизации запросов особо остро стоит в нагруженных веб-базируемых системах. В таких условиях оптимизация SQL-запросов выходит на первый план: скорость выполнения запросов определяет производительность всех вышестоящих процедур, и таким образом, оптимизация других компонентов системы лишена смысла, если не была предварительно проведена оптимизация используемых запросов.

Проблема оптимизации запросов требует решения следующих задач:

1. Преобразование запроса к более эффективному непроцедурному представлению (логическая оптимизация).
2. Выбор набора альтернативных процедурных планов выполнения запроса.
3. Оценка стоимости выполнения запроса по выбранному плану и т.д.

Основными ресурсами, которые расходуются при выполнении запроса, являются [1]:

1. Ресурсы процессора.
2. Ресурсы устройств внешней памяти.
3. Сетевые ресурсы.

Для распределённых запросов основную роль играют сетевые ресурсы, и, следовательно, для оптимизации таких запросов необходимо найти оптимальный глобальный план выполнения запроса.

В прошлой статье [2] была предложена графовая модель, к которой был применён муравьиный алгоритм для поиска оптимального плана выполнения запроса. Целью данной работы является исследование параметров муравьиного алгоритма, который используется для поиска оптимального пути по графу. Граф представляет собой все возможные глобальные планы выполнения распределённого запроса.

Общая постановка задачи

Муравьиный алгоритм в своей работе использует следующие параметры: коэффициент испарения феромона, коэффициент распыления феромона и количество муравьёв в колонии. Коэффициент распыления феромона предлагается вычислять динамически на основании весов рёбер графа. Два остальных параметра необходимо исследовать с целью определения их оптимального значения:

1. Количество муравьёв. Этот параметр определяет реальное количество итераций муравьиного алгоритма. Чем меньше его значение и чем сложнее исходный граф, тем менее точным будет результат. Однако слишком большое значение этого параметра может увеличить время выполнения алгоритма без улучшения результата. Количество вершин графа глобальных

планов выполнения запроса зависит от сложности запроса. Следовательно, количество муравьёв в колонии также должно вычисляться динамически на основании размера графа. Для вычисления количества муравьёв предлагается использовать линейную зависимость:

$$\text{AntsCount} = \text{VertexCount} * \text{AntsK},$$

где AntsCount – количество муравьёв; VertexCount – количество вершин графа; AntsK – коэффициент линейной зависимости.

- Коэффициент испарения феромона. Феромон – основной критерий муравьиного алгоритма, значение которого влияет на поиск оптимального решения. После прохождения каждой итерации алгоритма происходит испарение феромона. Коэффициент испарения определяет, как много феромона исчезнет с каждого ребра графа. Увеличение этого коэффициента может привести к увеличению хаотичности находимых результатов, в то время как уменьшение – к нахождению локального оптимума. Формула испарения феромона для каждого ребра графа:

$$F_{\text{new}} = F_{\text{old}} * (1 - \text{PheromoneElapseK}),$$

где F_{new} – количество феромона после испарения; F_{old} – старое количество феромона; PheromoneElapseK – коэффициент испарения феромона.

Задача данной работы – подобрать наиболее оптимальное соотношение этих параметров.

Исходные данные

Критерий останова

В качестве критерия останова алгоритма выбран критерий выполнения заданного количества итераций. Во-первых, он не требует дополнительных сравнений, что положительно сказывается на времени выполнения алгоритма; во-вторых, количество итераций алгоритма (количество муравьёв в колонии) находится в прямой зависимости от количества вершин графа.

Тестовый граф

Граф, использованный для исследования параметров алгоритма, представлен на рисунке 1. Названия вершин графа имеют формат <Порядковый номер>:<Название сервера распределённой БД>. Вес ребра графа определяет агрегированную стоимость соединения с сервером и извлечения данных. Он может зависеть от производительности сервера, от полосы пропускания сервера и её текущей загрузки, и т.д.

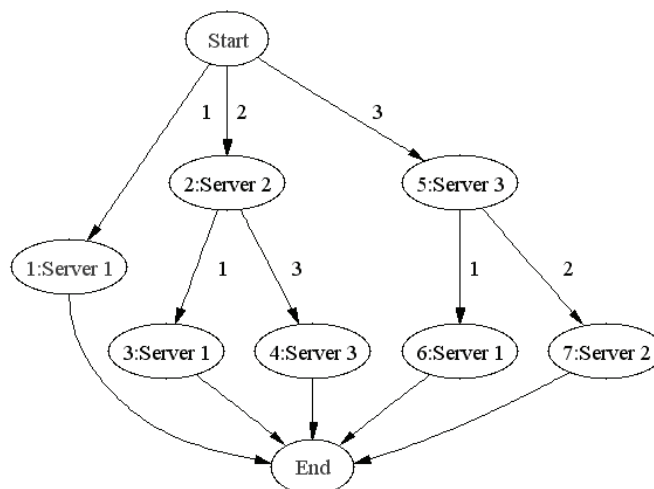


Рисунок 1. Тестовый граф

Граф обладает несколькими отличительными особенностями: разность между весами рёбер невелика, что приближает моделируемую задачу к реальной; в то же время существует одно очевидное наиболее оптимальное решение, что упрощает проверку результата: «Start» – «1:Server1» – «End».

Исследование параметров алгоритма

Суть исследования заключалась в том, чтобы выявить зависимость между значениями исследуемых параметров и точностью возвращаемого результата. Для этого алгоритм запускался многократно для каждого значения исследуемого параметра, и подсчитывалась доля правильных решений.

Коэффициент количества муравьёв

На рисунке 2 представлен график зависимости точности найденного решения от коэффициента количества муравьёв. Коэффициент испарения феромона при этом равен 0,1.

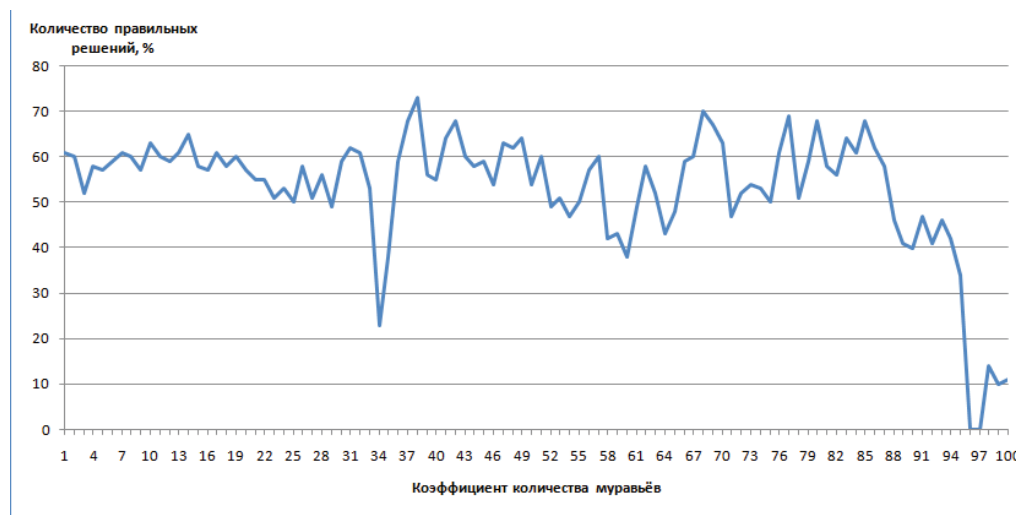


Рисунок 2. Зависимость точности решения от коэффициента количества муравьёв

Анализ результата позволяет опеределить оптимальное значение коэффициента – 38.

Коэффициент испарения феромона

На рисунке 3 представлен график точности найденного решения от коэффициента испарения феромона. Коэффициент количества муравьёв при этом равен 38.

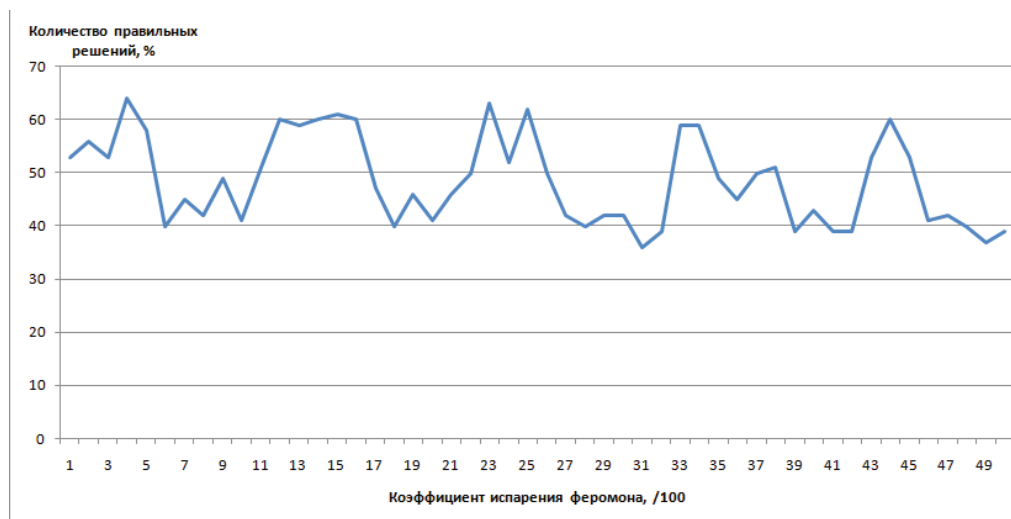


Рисунок 3. Зависимость точности решения от коэффициента испарения феромона. Лучшее значение коэффициента по графику = 0,04

Проверка найденных параметров

Для проверки двух найденных параметров был построен график зависимости количества правильных решений от коэффициента количества муравьёв, используя скорректированные параметры. График представлен на рисунке 4.

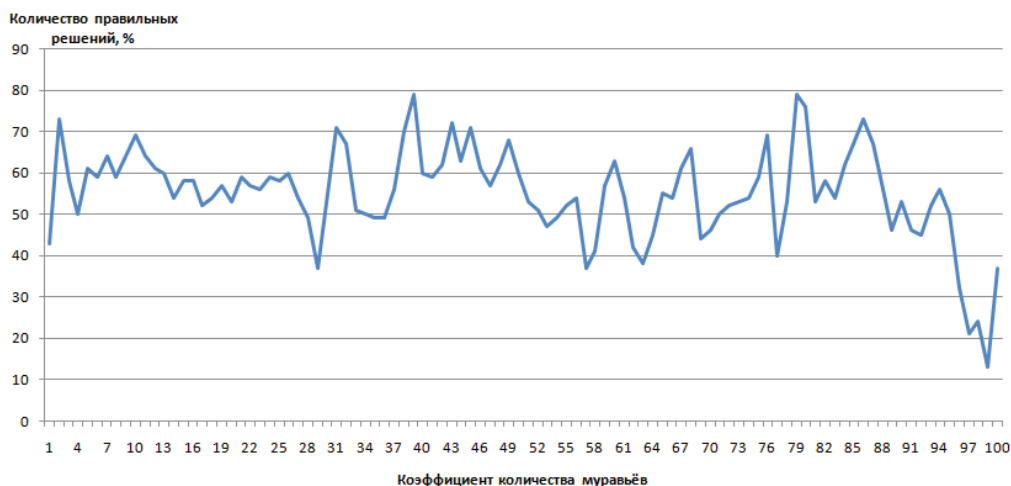


Рисунок 4. Проверка параметров

На этом графике наблюдается незначительное улучшение оптимального результата по сравнению с исходными параметрами: доля правильных результатов увеличилась на 6%.

Таким образом, в результате исследования получены следующие значения параметров муравьиного алгоритма:

Коэффициент количества муравьёв – 38;

Коэффициент испарения феромона – 0,04.

Полученные значения параметров проверены на тестовом графе путём запуска алгоритма поиска оптимального пути. Для 5000 итераций количество найденных оптимальных решений – 3793, что составляет примерно 75% от всех возможных решений.

Выводы

В ходе данной работы были исследованы параметры муравьиного алгоритма, который применяется для поиска оптимального глобального плана запроса.

Предварительно для этой цели были выбраны критерий останова и структура тестового графа для исследования параметров, которая соответствует необходимым для тестирования свойствам.

В результате были найдены параметры, при которых алгоритм работает наиболее оптимально:

Коэффициент количества муравьёв – 38;

Коэффициент испарения феромона – 0,04.

Точность результата работы алгоритма – 75%.

Литература

- [1] Матиас Ярке, Юрген Кох. Оптимизация запросов в системах баз данных // *Computing Surveys*, Vol. 16, No. 2, June 1984.
- [2] Заславский В., Савкова Е. Оптимизация выполнения распределённых запросов.