

Применение параллельных вычислений для повышения эффективности генетического поиска

Д.И. Батищев, Е.А. Неймарк, Н.В. Старостин

Нижегородский государственный университет им.Н.И.Лобачевского

Появление высокопроизводительных вычислительных машин позволяет исследователям решать за приемлемое время более сложные оптимизационные задачи, размер которых превышает 2^{100} . Эволюционные алгоритмы успешно применяются для решения оптимизационных задач различных уровней сложности, а использование современных компьютерных технологий повышает производительность поиска. Таким образом, сочетание эволюционных алгоритмов, использование высокопроизводительной техники и параллельных технологий является перспективной областью для решения оптимизационных задач. Данный класс алгоритмов называется параллельными или распределенными генетическими алгоритмами (parallel or distributed genetic algorithms). В отечественной литературе термин «параллельный генетический алгоритм» часто относится к мультипопуляционным моделям, без применения параллельных технологий вычислений. Применение параллельных технологий при решении задач оптимизации позволяет не только увеличить скорость поиска, а в некоторых случаях также увеличивает адаптивные свойства алгоритма. Рассмотрим основные направления параллельных генетических алгоритмов.

Обзор параллельных генетических алгоритмов

Можно отметить два основных направления: (1) применение параллельных технологий без изменения структуры генетического алгоритма; (2) изменение структуры генетического алгоритма, путем введение ограничения на возможность скрещивания и дальнейшее применение параллельных технологий. К первому направлению относится применение технологии master-slave. Ко второму направлению - мелкоструктурированный параллельный генетический алгоритм и крупноструктурированный параллельный генетический алгоритм.

Применение технологии master-slave. Первое направление связано с особенностями структуры генетического алгоритма. Элементарные операции, такие как построение особи, в операторе создания начальной популяции, построение потомка и его мутация, в операторе воспроизводства, а также оценка приспособленности особи, являются независимыми и легко поддаются параллельной обработке. Первые работы по распараллеливанию генетических алгоритмов велись в данном направлении. В данной группе алгоритмов работа над популяцией проводится на главном процессоре, а подчиненные процессоры вычисляют только приспособленность особей. То есть используется технология master-slave. При этом основная структура генетического алгоритма не меняется.

Ожидается, что при применении параллельных вычислений будет достигнуто повышение скорости нахождения квазиоптимального решения. Очевидно, что в данной схеме существенное влияние на скорость вычислений оказывает время, затрачиваемое на обмен данными между главным процессором и подчиненным, кроме того, продолжение работы возможно только после завершения вычислений всеми подчиненными процессорами. Таким образом при использовании данного метода, включение в работу большего числа процессоров может принести обратный эффект.

Крупноструктурированный параллельный генетический алгоритм. Вторая группа параллельных алгоритмов имеет свои корни в мультипопуляционных моделях, таких как метод случайных иммигрантов в исследованиях Grefenstette, метод ниш и островных моделях. В данном случае структура генетического алгоритма изменяется, популяция делится на группы (локальные популяции), развивающиеся параллельно и независимо

друг от друга, при этом они имеют возможность иногда обмениваться особями. При этом каждая группа обычно обрабатывается на отдельном процессоре. Основным отличием этих методов от классического генетического алгоритма является ограничение по скрещиванию особей: особи разных групп не могут образовывать брачную пару. Кроме того естественный отбор действует тоже только внутри одной группы.

За основу параллельного генетического алгоритма (ПГА) выбирается некоторый метод разбиения популяции, затем к нему применяются параллельные технологии.

После разбиения популяции на несколько групп, возникает вопрос о том насколько изолированными эти популяции должны быть. Таким образом, во-первых, должна быть определена структура связей между популяциями, а, во-вторых, какими данными и как часто они должны обмениваться.

Структура связей между популяцией в этой разновидности ПГА обычно отражает архитектуру связей между используемыми компьютерами или архитектуру связей между процессорами, чтобы минимизировать затраты на обмен данными. Тем не менее, организация связей между популяциями имеет большое влияние на эффективность поиска. При наличии большого числа связей информация распространяется на большее число групп, но это замедляет алгоритм в целом. При обратной ситуации может иметь место преждевременная сходимость в ряде локальных популяций, что понижает эффективность поиска. Обмен лучшими решениями осуществляется с двумя целями: (1) предотвращение преждевременной сходимости, для этого отслеживается разнообразие в каждой группе; (2) распространение лучшего найденного решения в группах для повышения общей приспособленности особей.

Алгоритмы данного типа широко применяются исследователями, поскольку: (1) принципиально не отличается от классического генетического алгоритма, добавляется только блок обмена решениями; (2) легко могут быть реализованы в сети на отдельных машинах и даже на однопроцессорных машинах.

Мелкоструктурированный параллельный генетический алгоритм. Третья группа параллельных алгоритмов имеет черты первых двух групп. Эти алгоритмы имеют одну популяцию, однако внутри популяции у особей существуют ограничения на скрещивание и конкуренцию. Популяция имеет регулярную структуру, согласно которой для каждой особи определяется окрестность, и только внутри нее для особи может быть выбрана пара или конкурент для естественного отбора. Таким образом, чем больше окрестность, тем ближе алгоритм к классическому генетическому алгоритму без ограничений. Структура популяции естественным образом проецируется на распределенную архитектуру: каждая особь ассоциируется с одним процессором. В силу такого распределения, в мелкоструктурированных ПГА обычно применяются поколенческие стратегии естественного отбора с полной заменой родителей.

Такой алгоритм является более гибким, чем крупноструктурированный, поскольку окрестность может меняться в ходе работы, кроме того в силу перекрытия окрестностей происходит распространение лучшей особи по всей популяции.

Заключение

В представленном обзоре, приводятся основные группы параллельных генетических алгоритмов. Кроме того, исследователями часто используются гибридные ПГА, где в локальных популяциях крупноструктурированного ПГА применяются технологии master-slave для вычисления приспособленности или мелкоструктурный ПГА и т.п.

Поскольку генетические алгоритмы наиболее успешно применяются на задачах с большим объемом вычислений, то для них актуально применение параллельных технологий для повышения скорости вычислений, а следовательно и эффективности поиска.

Тем не менее, до сих пор мало исследованы теоретические оценки эффективности применения ПГА по сравнению с последовательным генетическим алгоритмом.