

СПИСОК ЦИТИРОВАННЫХ ИСТОЧНИКОВ

1. Головкин, В.А. Нейроинтеллект: теория и применение. Книга 1: Организация и обучение нейронных сетей с прямыми и обратными связями. - Брест: БПИ, 1999.

2. Оленев, Н.Н. Основы параллельного программирования в системе MPI. - М.: ВЦ РАН, 2005.

Материал поступил в редакцию 25.09.2008

SAVITSKY Yu.V. The comparative analysis of power of neural network functioning parallel schemes

In this article the results of power of parallel neural network functioning algorithm are considered. The parallel schemes are made as program system based on the MPI (Message Passing Interface) protocol. The computational experiments of time dependence from number of neurons, number of training patterns are discussed. The analysis of parallel algorithms time complexity with MPI message passing and without MPI message passing is given.

УДК 681.324

Савицкий Ю.В.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ПРОИЗВОДИТЕЛЬНОСТИ СХЕМ ПАРАЛЛЕЛИЗАЦИИ АЛГОРИТМОВ ОБУЧЕНИЯ НЕЙРОННОЙ СЕТИ

Введение. Как показывает практика, в нейросетевых методах обработки информации наиболее трудоемким этапом в вычислительном и временном аспекте является процесс обучения нейронной сети [1]. Процесс обучения, также как и процесс функционирования, характеризуются параллельной архитектурой и соответственно возможностью параллелизации и увеличения производительности.

В статье автора «Сравнительный анализ производительности схем параллелизации обучения нейронной сети» были рассмотрены схемы параллелизации функционирования многослойного персептрона "один нейрон для одного вычислительного модуля (ВМ)", "группа нейронов для одного ВМ", "слой за слоем", "использование копий нейронной сети".

Задачей данной работы является всесторонне исследование производительности вышеуказанных схем параллелизации, примененных в задаче обучения многослойной нейронной сети.

Сравнительный анализ схем параллелизации. При обучении нейронной сети кроме вычисления выходных значений нейронных элементов всех слоев (функционирование нейронной сети) производится вычисление значений ошибок нейронных элементов, изменение весовых коэффициентов и порогов нейронных элементов каждого слоя и для каждого эталона. Следовательно, это более длительный и сложный в вычислительном аспекте процесс, чем функционирование нейронной сети.

Производительность параллельных алгоритмов обучения нейронной сети, будем оценивать при помощи временных соотношений их исполнения на базе разработанной MPI-программы [2]. Чем меньше время выполнения всех действий для некоторого числа эталонов нейронной сети, тем выше производительность алгоритма обучения (количество "обработанных" эталонов за определенное время).

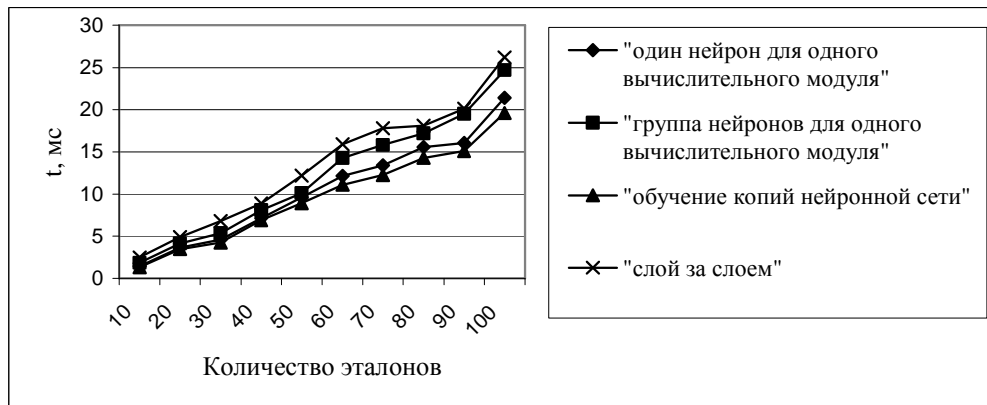


Рис. 1. График временной зависимости обучения нейронной сети для параллельных алгоритмов от количества эталонов без учета межмодульных пересылок

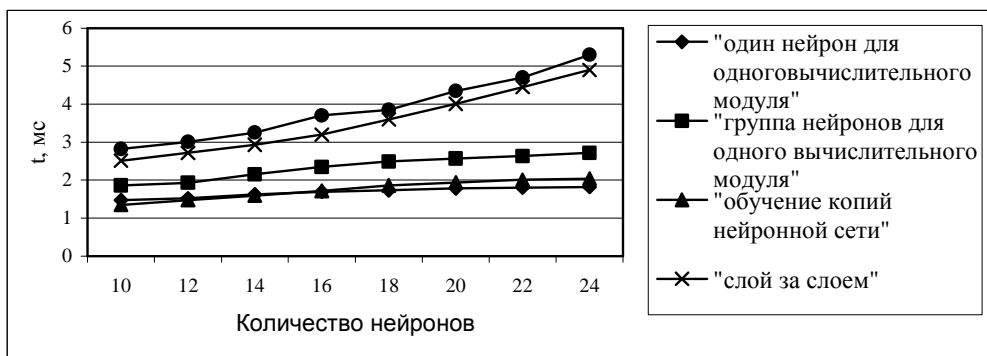


Рис. 2. График временной зависимости обучения нейронной сети для параллельных алгоритмов от количества нейронов в слоях без учета межмодульных пересылок

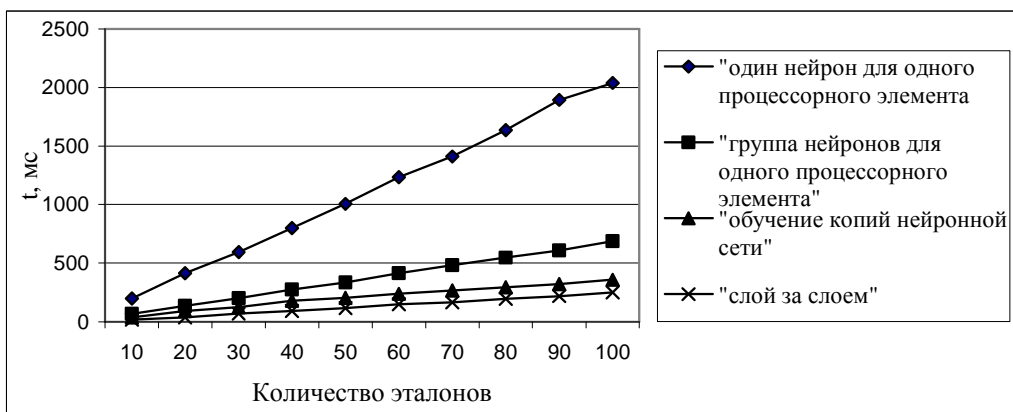


Рис. 3. График временной зависимости обучения нейронной сети для параллельных алгоритмов от количества эталонов с учетом межмодульных пересылок

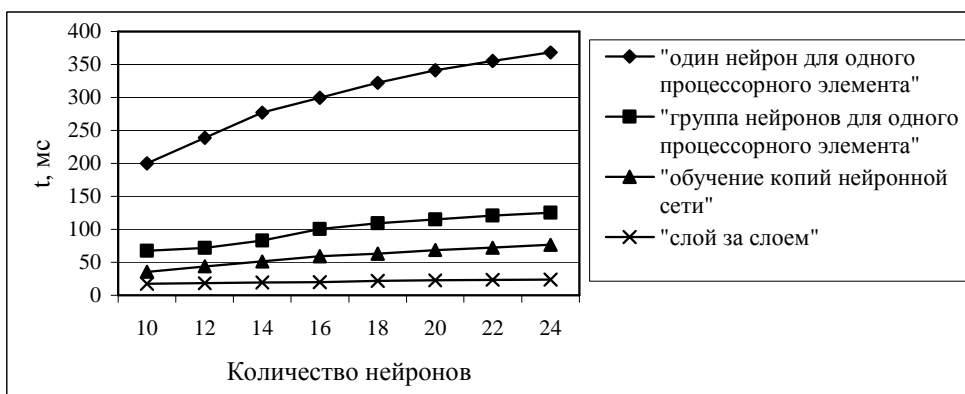


Рис. 4. График временной зависимости обучения нейронной сети для параллельных алгоритмов от количества нейронов в слоях с учетом межмодульных пересылок

Сначала проанализируем временные соотношения обучения нейронной сети без учета времени межмодульных пересылок в зависимости от количества обучающих эталонов и в зависимости от количества нейронных элементов для всех параллельных алгоритмов обучения.

Для этого проведем ряд тестов, в которых будем использовать разное количество входных эталонов и разное количество нейронных элементов во входном и скрытом слоях. Результаты тестов отображены на рис. 1 и рис. 2.

Из графиков видно, что самое большое время обучения без учета межмодульных пересылок получается при использовании алгоритма "слой за слоем" и последовательного алгоритма, так как в этом случае каждый процесс "обрабатывает" все нейроны слоя. При использовании алгоритмов, в которых каждый процесс производит вычислительные операции над меньшим числом нейронных элементов, время функционирования нейронной сети получается меньше. В алгоритме "обучение копий нейронной сети" небольшое время обучения достигается за счет того, что одновременно производятся вычислительные операции над двумя эталонами. Время функционирования нейронной сети в большей степени зависит от количества входных эталонов, чем от количества нейронов в слоях. Увеличение числа нейронов в слоях практически не влияет на время обучения нейронной сети при использовании алгоритма "один нейрон для одного вычислительного модуля", так как все процессы выполняют вычислительные операции над одним нейроном, но при этом увеличивается число создаваемых процессов.

Из графиков видно, что самое большое время обучения без учета межмодульных пересылок получается при использовании алгоритма "слой за слоем" алгоритма, так как в этом случае каждый процесс "обрабатывает" все нейроны слоя. При использовании алгоритмов, в которых каждый процесс производит вычислительные

операции над меньшим числом нейронных элементов, время функционирования нейронной сети получается меньше. В алгоритме "обучение копий нейронной сети" небольшое время обучения достигается за счет того, что одновременно производятся вычислительные операции над более чем одним эталонами. Время функционирования нейронной сети в большей степени зависит от количества входных эталонов, чем от количества нейронов в слоях. Увеличение числа нейронов в слоях практически не влияет на время обучения нейронной сети при использовании алгоритма "один нейрон для одного вычислительного модуля", так как все процессы выполняют вычислительные операции над одним нейроном, но при этом увеличивается число создаваемых процессов.

Также проведем ряд тестов для оценки времени обучения нейронной сети в зависимости от количества входных эталонов и в зависимости от количества нейронных элементов в слоях, но с учетом межмодульных пересылок. Их результаты отображены на рис. 3 и рис. 4. Из графиков видно, что время обучения многослойного персептрона возросло в несколько раз. Также видно, что время обучения нейронной сети получается меньше, если использовать алгоритмы, в которых используется меньшее число обменов информацией между процессами ("слой за слоем" и "обучение копий нейронной сети"). То есть более производительными получаются те алгоритмы обучения нейронной сети, в которых используется меньше межпроцессных обменов. Самым непроизводительным является алгоритм "один нейрон для одного вычислительного модуля", из-за использования большого числа обменов данными между процессами.

Выводы. Таким образом, производительность параллельных схем алгоритмов обучения можно повысить: 1) за счет использования вычислительных систем с меньшим временем MPI-обменов; 2) за счет минимизации в алгоритмах количества MPI-обменов 3) при увеличении вычислительной сложности MPI-процессов.

СПИСОК ЦИТИРОВАННЫХ ИСТОЧНИКОВ

1. Головкин, В.А. Нейроинтеллект: теория и применение. Книга 1: Организация и обучение нейронных сетей с прямыми и обратными связями. - Брест: БПИ, 1999.

2. Оленев, Н.Н. Основы параллельного программирования в системе MPI. - М.: ВЦ РАН, 2005.

Материал поступил в редакцию 25.09.2008

SAVITSKY Yu.V. The comparative analysis of power of neural network training algorithms parallel schemes

In this article the results of power of parallel neural network training algorithms are considered. The parallel schemes are made as program system based on the MPI (Message Passing Interface) protocol. The computational experiments of time dependence from number of neurons, number of training patterns are discussed. The analysis of parallel algorithms time complexity with MPI message passing and without MPI message passing is given.

УДК 519.876.5: 544.77.023.5

Дереченник А.С., Дереченник С.С.

ОСОБЕННОСТИ ПЕРКОЛЯЦИИ В НЕРЕГУЛЯРНЫХ ДИСПЕРСНЫХ СТРУКТУРАХ

Введение. Важные прикладные свойства композиционных материалов (композитов) – влаго- и газопроницаемость – обусловлены особенностями структуры их капиллярно-пористого пространства, которая является в существенной степени неупорядоченной. Общая ситуация, связанная со случайным распространением жидкости через среду, когда термины «жидкость» и «среда» интерпретируются в соответствии с физическим смыслом задачи, рассматривается теорией протекания, или перколяции [1, 2]. Применительно к исследованию влагопроницаемости композитов, жидкостью, очевидно, является вода, а под средой понимается дисперсная структура, т.е. система с порами в качестве частиц дисперсной фазы.

Важнейшим понятием теории протекания является порог перколяции – значение критической концентрации, например, пористой фазы, при которой система становится проницаемой для жидкости.

Наиболее известные задачи перколяции на регулярных (квадратных, треугольных, кубических и т.п.) решетках исследованы достаточно полно, для них доказан ряд строгих утверждений и соотношений, поэтому на сегодняшний день такие задачи представляют интерес лишь с теоретической точки зрения. Поскольку в реальном материале поры разновелики и случайно размещены в объеме, структура его порового пространства однозначно нерегулярна и не может быть адекватно представлена на регулярных решетках. Закономерности же протекания в нерегулярных (неупорядоченных, полидисперсных) структурах до настоящего времени остаются малоизученными.

В настоящей работе предложена методика идентификации перколяционных кластеров, а также исследованы некоторые особенности перколяции в нерегулярных дисперсных структурах, являющихся моделями порового пространства композиционных материалов.

Методика идентификации перколяционных кластеров в нерегулярной структуре. Для анализа используются неупорядоченные многочастичные моно- и полидисперсные системы, синтезированные по разработанному ранее алгоритму и представленные в виде стохастической матрицы связности элементов структуры [3]. Исходная структура, в двумерной модельной области (контейнере квадратной формы) представляет собой случайную упаковку n частиц в виде круглых дисков одинакового или различного размера. Кластером считается такая группа частиц, в которой для каждой частицы можно было бы найти хотя бы одну, контактирующую с ней и также принадлежащую группе. Определение частиц или стенок контейнера, контактирующих с любой заданной частицей, выполняется по матрице связности и сводится к поиску единичных элементов матрицы.

Коэффициент плотности упаковки исходной системы достаточно высок (на уровне $\eta = 0.80$), поэтому все частицы принадлежат одному кластеру. Разработанная методика основывается на удалении из исходной структуры некоторого количества частиц, после чего оставшиеся частицы могут образовывать два и более кластеров меньшего размера, вплоть до одиночных (изолированных) частиц.

Процесс идентификации перколяционного кластера включает следующие действия:

- 1) Согласно задаваемой относительной доле p , определяется количество дисперсных частиц $m = pn$, которые будут помечены (окрашены), составляя, таким образом, поровое пространство модельного образца. Если система полидисперсная, заданная доля применяется к частицам каждого типа (размера) в наборе $n = \sum n_i$, т.е. $m = \sum (m_i = pn_i)$.

- 2) Определяются конкретные частицы, подлежащие окраске. Они могут выбираться случайно и независимо (равновероятно), либо задаваться в соответствии с определенными правилами. Первый вариант моделирует случайное протекание, или перколяцию Бернулли, второй – коррелированную перколяцию (в данной работе исследуется случайная перколяция). Строки и столбцы матрицы связности, соответствующие оставшимся $(n - m)$ частицам, вычеркиваются из матрицы.

- 3) Выполняется поиск частиц, контактирующих с одной из стенок контейнера (например, левой либо верхней). Найденные частицы составляют множество точек возможной инъекции жидкости. Если множество окажется пустым, то инъекция неосуществима, значит, невозможно и протекание жидкости через образец в данном направлении (слева либо сверху). Следует вернуться к п. 2 для генерации нового варианта окраски частиц.

- 4) Одна из частиц найденного множества определяется как центр инъекции, после чего выполняется поиск частиц, с ним контактирующих, затем – новых частиц, контактирующих с каждой из найденных и т.д. Каждая найденная частица присоединяется к растущему таким образом кластеру (например, перекрашивается в новый цвет). Процесс продолжается до тех пор, пока возможно присоединение к кластеру еще хотя бы одной новой частицы. Если после этого хотя бы одна из перекрашенных частиц контактирует с правой (либо нижней) стенкой, кластер идентифицируется как перколяционный, т.к. возможно протекание жидкости через область моделирования слева направо (либо сверху вниз). Примеры таких кластеров представлены на рисунке 1. В противном случае найденный кластер считается кластером конечного размера, а п. 4 повторяется для другого центра инъекции. Если ни для одной из частиц – центров инъекции перколяционный кластер не получен, выполняется генерация нового варианта окраски согласно п. 2.

Предложенная методика применена и к менее плотным, в том числе разреженным структурам, в которых большинство частиц не находятся в непосредственном контакте друг с другом (например, сравнительно крупные поры некоторого материала могут соединяться небольшими капиллярными каналами). При формировании матрицы связности такой структуры задается максимальное расстояние, при котором частицы еще считаются связными.

Дереченник Станислав Станиславович, к.т.н., доцент, заведующий кафедрой «ЭВМ и системы» Брестского государственного технического университета.

Беларусь, БрГТУ, 224017, г. Брест, ул. Московская, 267.

Физика, математика, информатика