

УДК 519.71

О.В. ЩЕРБА

*Черниговский государственный технологический университет, Украина***ПРИМЕНЕНИЕ УСТРОЙСТВА СОГЛАСОВАНИЯ НАГРУЗОК
ДЛЯ ПОВЫШЕНИЯ НАДЕЖНОСТИ ФУНКЦИОНИРОВАНИЯ
РАСПРЕДЕЛЕННОЙ КОМПЬЮТЕРНОЙ СИСТЕМЫ**

Рассмотрен нейросетевой подход к построению устройства согласования нагрузок распределенной компьютерной системы и задача аппаратной реализации данного устройства. Указаны особенности применения нейронных сетей в задаче распределения нагрузки между узлами системы, а также достоинства аппаратной реализации.

распределенная компьютерная система, искусственная нейронная сеть, ПЛИС, прогнозирование**Введение**

Одной из основных задач при создании распределенных вычислительных систем, является задача управления распределением ресурсов, находящихся в системе. Качество решения задачи управления ресурсами напрямую определяет степень надежности функционирования распределенной системы. Неправильное распределение ресурсов может привести к тупиковым ситуациям в работе системы, или же к простоя части ресурсов. Исходя из этого определяется актуальность создания средств оптимального управления ресурсами вычислительной системы с целью повышения надежности ее функционирования.

Задачу управления и распределения ресурсов предложено решать путем включения в распределенную систему устройства согласования нагрузок (диспетчера), которое выполняет централизованный учет и управление ресурсами.

Алгоритмический подход к решению задачи реализации диспетчера является неэффективным. Это связано с отсутствием точных алгоритмов оценивания потребности выполняющихся задач в распределенной системе в ресурсах, а так же выбора ресурсов для выполнения. Существующие решения позволяют произвести лишь приблизительную оценку

ресурсов, не гарантируя тем самым оптимальность распределения задач между ресурсами. Таким образом, можно утверждать, что процесс оценивания ресурсов является слабо формализованным [1]. Решать такого рода задачи призваны нейронные сети. Проведя анализ, было принято решение о применении нейросетевых технологий при построении диспетчера распределенной системы. В этом случае, замена алгоритмического подхода нейросетевым позволит исключить возможность тупиковых ситуаций, а следовательно, повысить отказоустойчивость системы.

С целью повышения быстродействия диспетчера, а также для снижения вычислительной нагрузки на основной процессор вычислительного узла распределенной системы, предложено создать диспетчер как внешнее устройство, реализованное на программируемых логических интегральных схемах (ПЛИС) и подключаемое к системе посредством PCI-шины вычислительного узла.

Перспектива применения ПЛИС заключается в том, что в пределах одной микросхемы можно реализовать большое количество параллельно функционирующих вычислительных узлов. Такой подход к реализации позволит получить максимальную производительность искусственной нейронной сети

(ИНС) за счет высокой степени параллелизма функционирования [4]. Одним из недостатков реализации ИНС в программируемой логической интегральной схеме является сложность ее обучения. Поэтому обучение сети целесообразно проводить на ее модели построенной в среде моделирования, а по результатам моделирования получить набор значений весовых коэффициентов, которые следует сохранить в ПЛИС.

Результаты исследования

Общая структура диспетчера приведена на рис. 1.

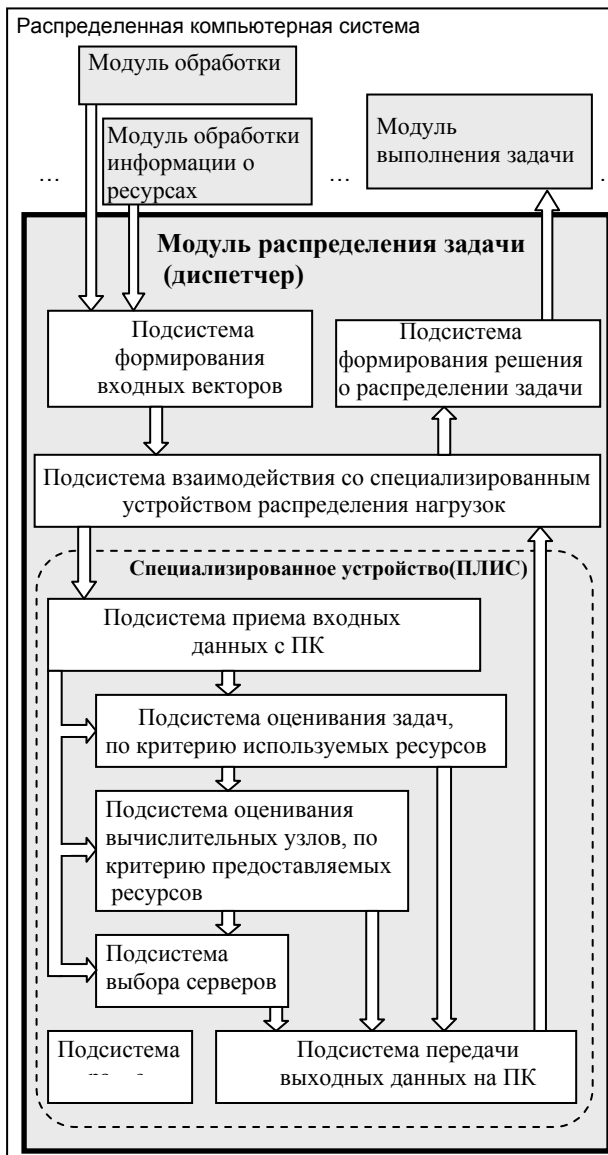


Рис. 1. Структура устройства согласования нагрузок (диспетчера) распределенной компьютерной системы

Как видно из рис. 1, диспетчер, состоит из трех основных подсистем, определяющих его целевые возможности [1]:

- подсистема оценивания задачи по критерию используемых ресурсов;
- подсистема оценивания узлов по критерию предоставляемых ресурсов;
- подсистема выбора серверов.

Подсистему оценивания задачи по критерию используемых ресурсов наиболее эффективно реализовать с помощью многослойного персептрона [2] со следующими параметрами:

- количество слоев – 3;
- количество нейронов входного слоя – 5;
- число нейронов скрытого слоя – 5;
- число нейронов выходного слоя – 4;
- активационная функция – сигмоидальная.

Подсистему оценивания узла по критерию предоставляемых ресурсов наиболее эффективно реализовать с помощью многослойного персептрона (отличительной особенностью которого являются то, что в выходном слое будет 1 нейрон) со следующими параметрами:

- количество слоев – 2;
- количество нейронов входного слоя – 8;
- число нейронов выходного слоя – 1;
- активационная функция – сигмоидальная.

Процесс выбора оптимальной комбинации вычислительных узлов в подсистеме выбора серверов будет сводиться к поиску минимального расстояния от точки оптимального распределения до точки с координатами (количество серверов, суммарный коэффициент комбинации) [1]. Для этого будет использоваться радиально-базисная сеть [2]. Главной особенностью этой сети является ее обучение. Процесс обучения радиальной сети сводится к подбору параметров радиальных функций. В данной ситуации приемлем наиболее общий случай – использование случайных параметров. Данный подход не гарантирует полной оптимальности решения, но

гарантирует существование решения.

Наибольшую вычислительную нагрузку при реализации искусственных нейронов создает реализация вычислений активационной функции, так как вычисление сигмоидальной функции является не тривиальной задачей для ПЛИС, однако использование этой функции оправдано функциональными требованиями к нейронной сети.

Вычисление сложных функций реализуется путем построения таблицы перекодировки (look-up tables) изображенной на рис. 2, которая ставит в соответствие входным значениям функции выходные.

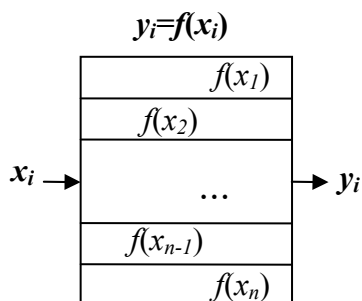


Рис. 2. Таблица перекодировки (look-up tables) для вычисления сложных математических функций

Данный способ, построения вычислений специфических функций является наиболее подходящим по критерию быстродействия, так как сводит вычисление функции обращением к ячейкам встроенных блоков памяти. Размер таблицы прямопропорционален требуемой точности вычислений. Построение таблицы выполняется перед началом работы устройства.

Решение возможно несколькими способами:

- реализация вычислителя активационной функции в виде отдельного элемента, к которому будут обращаться по очереди все нейроны;
- в каждом нейроне реализовать блок вычисления этой функции.

Первый способ, представленный на рис. 3, является подходящим с точки зрения снижения ресурсоемкости (сокращения количества ячеек ПЛИС требуемых для реализации нейросети), однако последовательное использование вычислительного блока

всеми нейронами значительно ограничивает эффективность параллельной работы нейронов и, следовательно, снижает производительность вычислений.

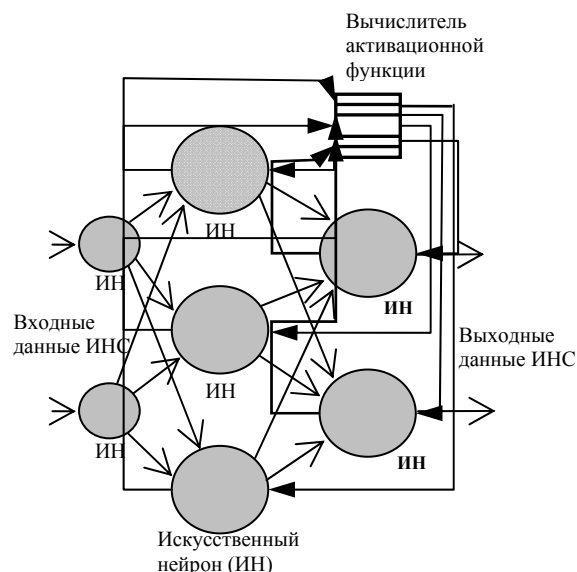


Рис. 3. Применение одного вычислителя активационной функции для всех нейронов сети

Второй способ, представленный на рис. 4 удовлетворяет требованию достижения максимального быстродействия, а также обеспечивает независимое функционирование нейронов.

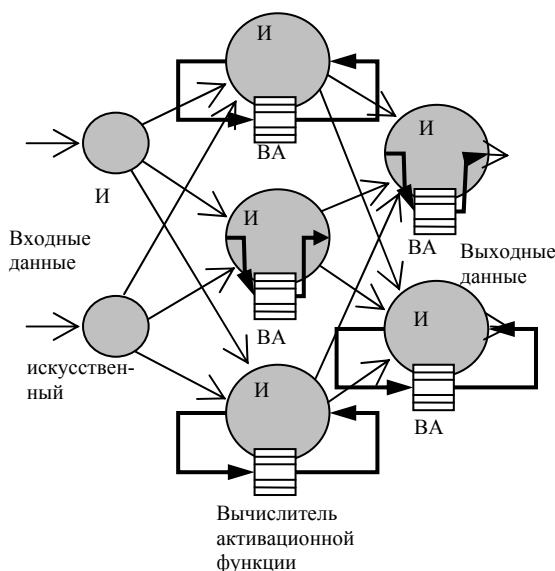


Рис. 4. Применение для каждого нейрона отдельного вычислителя активационной функции

Однако, такой подход к реализации является ресурсоемким, причем с возрастанием количества нейронов в сети возрастает потребность в ресурсах

для вычислителей активационной функции.

Так как сеть является многослойной, а работа слоев сети последовательна, то компромиссным решением может служить построение блоков вычисления сигмоидальной активационной функции, необходимой для независимой работы одного слоя с наибольшим количеством нейронов [3]. При дальнейшей работе нейроны каждого слоя будут обращаться к одному и тому же блоку вычислений. Такой подход, представленный на рис. 5, является самым подходящим в рассматриваемом варианте реализации ИНС, так как позволяет сократить требования к ресурсам без потери производительности сети в целом.

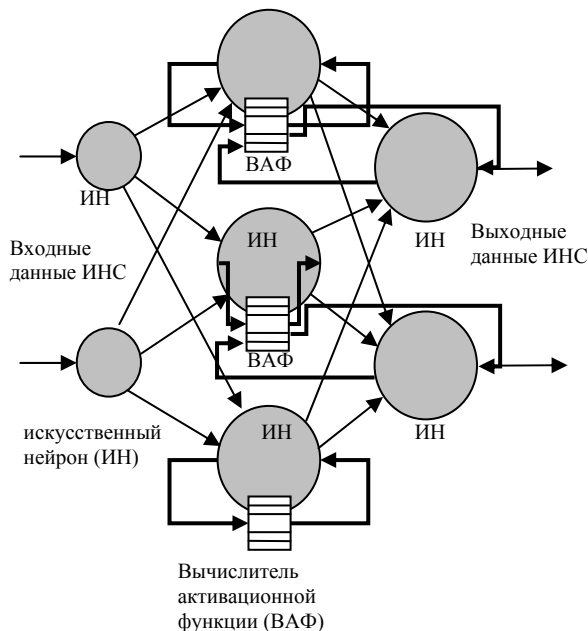


Рис. 5. Применение одного вычислителя активационной функции для нескольких нейронов из разных слоев нейронной сети

Выводы

Таким образом, в данной статье были рассмотрены особенности аппаратной реализации диспетчера распределенной компьютерной системы. Применение нейросетевых технологий позволит повысить качество управления ресурсами, а, следовательно, повысить надежность функционирования распределенной системы и избежать тупиковых ситуаций в работе системы, которые свойственны алгоритмическому подходу к диспетчеризации.

Предложено решение проблемы аппаратного вычисления активационной функции нейронов, которое заключается в применении таблиц перекодировки. С одной стороны, такой подход позволит повысить быстродействие за счет обеспечения параллельной работы нейронов одного слоя. С другой стороны, предложенный подход к реализации вычислителя активационной функции позволит сократить количество ресурсов за счет использования одних и тех же вычислителей нейронами из разных слоев.

Литература

1. Rudenko O., Gorelova O., Zaets A. The neural network coordination loading device for the distributed computer system // Комп'ютерні науки та інженерія: Матеріали 2-ї Міжн. конференції молодих науковців CSE-2007. – Львів: Видавництво Національного університету «Львівська політехніка». – 2007. – 232 с.
2. Руденко О.Г., Бодянский Е.В. Искусственные нейронные сети: Учебное пособие. – Х.: ООО «Компания СМИТ», 2005. – 408 с.
3. Руденко О.Г., Горелова О.В., Заец О.В. Устройство согласования нагрузок в распределенной компьютерной системе на основе нейронных сетей // Материалы 3-й Межд. научно-технической конференции молодых ученых и студентов «Информатика и компьютерные технологии 2007». – Донецк. – С. 124.
4. Руденко О.Г., Заец А.В., Горлова О.В., Пискун А.С. Устройство согласования нагрузок в распределенной компьютерной системе на основе нейронных сетей // Информационные технологии и информационная безопасность в науке, технике и образовании "ИНФОТЕХ - 2007": Материалы междунаучно-практической конференции. Часть 2. – Севастополь. – С. 83.

Поступила в редакцию 4.02.2008

Рецензент: д-р техн. наук, проф. Б.М. Конорев, Национальный аэрокосмический университет им. Н. Е. Жуковского «ХАИ», Харьков.