



УДК 681.5.13

**П.И. Кравец**, канд. техн. наук,  
**В.Н. Шимкович, Д.А. Ференс**  
Национальный технический университет Украины  
«Киевский политехнический институт»  
(Украина, 03056, Киев, пр-т Победы, д. 37,  
тел.: 4068346, e-mail: kravets@ntu-kpi.kpi.ua)

### **Метод и алгоритмы реализации на ПЛИС функции активации для искусственных нейронных сетей**

Предложены метод и алгоритм аппаратной реализации нейронных сетей на ПЛИС типа FPGA с использованием языка VHDL. Приведены примеры аппаратной реализации средствами ПЛИС искусственных нейронов и нейронных сетей с сигмоидальными функциями активации. Показано, что с помощью разработанного метода обеспечивается значительная скорость обработки информации и оптимизируется использование ресурса ПЛИС.

Запропоновано метод та алгоритм апаратної реалізації нейронних мереж на ПЛИС типу FPGA з використанням мови VHDL. Наведено приклади апаратної реалізації засобами ПЛИС штучних нейронів та нейронних мереж з сигмоїдальними функціями активації. Показано, що за допомогою розробленого методу забезпечується значна швидкість обробки інформації та оптимізується використання ресурсу ПЛИС.

*К л ю ч е в ы е с л о в а:* нейронные сети, функции активации, ПЛИС, язык VHDL.

Развитие теории автоматического управления связано с усложнением решаемых задач и повышением качественных показателей необходимых решений [1]. Традиционные методы управления основаны на теории линейных систем, в то время как реальные объекты являются, по своей природе, нелинейными. Нейросетевые системы управления представляют собой новое высокотехнологичное направление в теории управления и относятся к классу нелинейных динамических систем [2]. Высокое быстродействие посредством распараллеливания входной информации в сочетании со способностью к обучению нейронных сетей делает эту технологию весьма привлекательной для создания устройств управления в автоматических системах.

© П.И. Кравец, В.Н. Шимкович, Д.А. Ференс, 2015

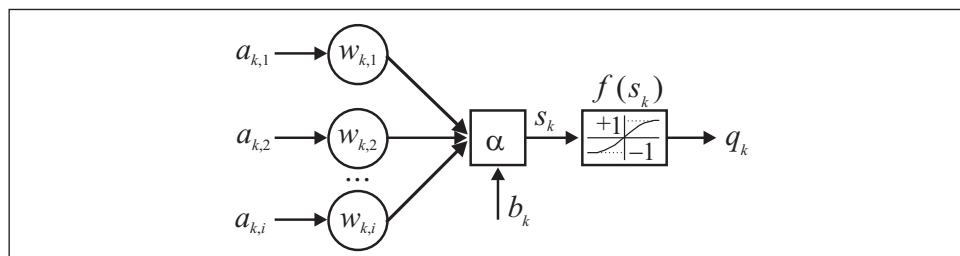


Рис. 1. Схема искусственного нейрона

В настоящее время основным методом реализации нейросетевых систем управления является программный, с использованием компьютерной техники или специализированных контроллеров, построенных на их основе, что значительно сокращает практическое применение нейросетевых систем управления. Высокая стоимость таких регуляторов делает их практически недоступными для использования в простых системах управления. Кроме того, компьютерные нейросетевые регуляторы имеют ограниченную производительность и требуют значительных затрат времени на обучение. Рекуррентность и последовательность действий процедуры обучения нейросети при реализации на всем множестве настроечных параметров не позволяет полностью решить проблему быстрого действия процедуры обучения нейронных сетей в одном темпе с динамикой объекта управления. Альтернативой этому является распараллеливание процедуры обучения и работы внутренних элементов нейронных сетей. Такая возможность появляется при аппаратной реализации нейронных сетей, построенных на нейрочипах или программируемых логических интегральных схемах (ПЛИС) [3].

Известны современные разработки, выполненные на ПЛИС высокой интеграции,— «нейрочип-8», инструментальная плата XDSP-680 на базе FPGA семейства Spartan компании Xilinx с нейросетевой прошивкой и другие перспективные разработки, такие как «нейрочип-2000» на основе FPGA Virtex / Virtex-E, а также набор инструментальных плат на основе ПЛИС различных серий и модулей различного назначения, позволяющие быстро и эффективно создавать вычислительные системы различного функционального назначения [4]. Все указанные разработки основаны на нейронах с пороговой функцией активации и предназначены для вычислительных систем общего назначения. Однако представляет интерес разработка нейроконтроллеров систем управления, способных функционировать и адаптироваться в реальном времени [7—12].

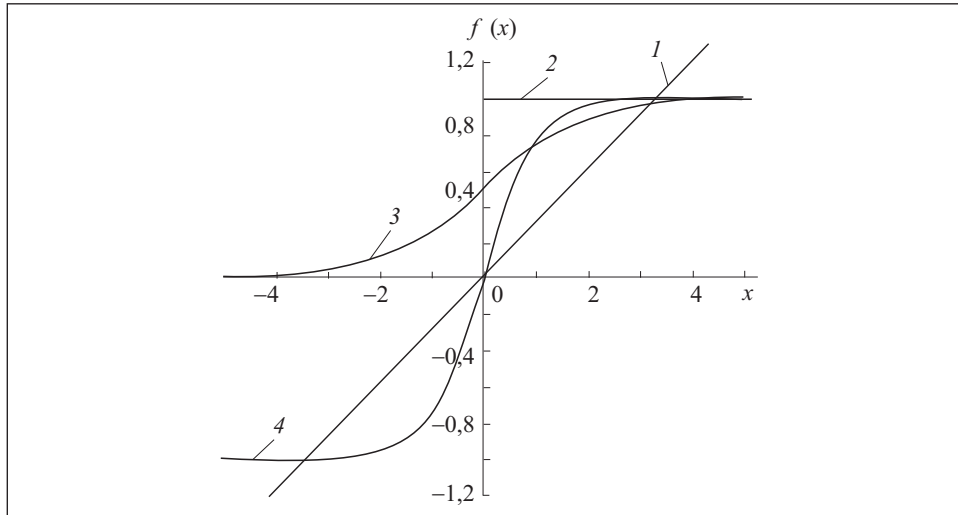


Рис. 2. Функции активации искусственных нейронов: 1 — линейная; 2 — пороговая; 3 — сигмоидальная; 4 — гиперболический тангенс

Предлагается метод и алгоритм аппаратной реализации искусственных нейронных сетей с функциями активации сигмоидального типа на программируемых интегральных логических схемах с использованием языка VHDL для дальнейшего проектирования на их основе компонентов нейросетевых систем управления, которые могут функционировать и адаптироваться в реальном времени.

**Результаты исследований.** Одна из проблем, возникающих при аппаратной реализации нейронных сетей, — реализация искусственного нейрона и его активационных функций средствами ПЛИС. На рис. 1 представлена структурная схема искусственного нейрона, где  $f$  — активационная функция нейрона;  $q_k$  и  $a_{k,i}$  — выходной сигнал и входящие сигналы  $k$ -го нейрона;  $w_{k,i}$  и  $b_k$  — синаптический вес и смещения  $k$ -го нейрона.

В искусственных нейронных сетях наибольшее распространение (по аналогии с биологическими процессами) получили функции активации, такие как линейная, пороговая и сигмоидальная (рис. 2).

Сигмоидальные функции — это наиболее широко используемый тип активационных функций. Они являются монотонно возрастающими, непрерывными и дифференцируемыми. Это класс функций, которые описываются выражением

$$f(x, k, b, T, c) = k + \frac{c}{1 + be^{Tx}}, \quad (1)$$

где  $x, k, b, T, c$  — параметры  $k \in R; b \in R, b > 0; T, c \in R \setminus \{0\}$ . Если принять  $k = 0, c = 1, b = 1$  и  $T = -1$ , то выражение (1) примет вид

$$f(x, 0, 1, -1, 1) = 0 + \frac{1}{1 + e^{-1x}} = \frac{1}{1 + e^{-x}}. \quad (2)$$

Функция, описанная выражением (2), называется «классический» сигмоид (см. рис. 2, 3). Если принять  $k = 1, c = -2, b = 1$  и  $T = 2$ , то выражение (2) примет вид

$$f(x, 1, 1, 2, -2) = 1 + \frac{-2}{1 + e^{2x}} = \frac{1 + e^{2x} - 2}{1 + e^{2x}} = \frac{e^{2x} - 1}{e^{2x} + 1} = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \quad (3)$$

Функция, описанная выражением (3), называется гиперболический тангенс (см. рис. 2, 4).

Сложными относительно цифровой реализации являются сигмоидальные функции активации. При разработке метода и алгоритма реализации нейронов с сигмоидальной функцией активации следует учитывать особенности строения ПЛИС, имеющей ресурсы поисковой таблицы, мультиплексоры, сумматоры, блоки умножения, но не имеющей блоков деления. Для нейрона необходима регулируемая точность вычислений и разрядность входных данных.

Для цифровой реализации нелинейных функций используют различные методы аппроксимации, такие как табличный, разложение в ряд Тейлора, кусочно-линейная аппроксимация и др. Метод разложения в ряд Тейлора требует выполнения многих умножений, а потому неприемлем для реализации в ПЛИС [4, 5]. Табличный метод предусматривает создание глобальной переменной, таблицы возможных значений целевой функции, непредсказуемый и неконтролируемый доступ к ней всеми нейронами сети, что, в свою очередь, создает большую временную задержку. Создание отдельной локальной таблицы для каждого нейрона в сети неприемлемо при использовании ресурса ПЛИС. Поэтому наиболее оптимальным методом реализации функций активации сигмоидального типа является кусочно-линейная аппроксимация.

Рассмотрим следующий метод реализации искусственного нейрона с функциями активации сигмоидального типа. Входными данными при этом являются функции, описываемые формулой (1).

Э т а п 1. Исследуем функцию активации на симметричность относительно осей. Запишем функцию (2) в виде

$$f(-x) = 1 - f(x) = 1 - \frac{1}{1 + e^{-x}} = 1 - \frac{1}{1 + \frac{1}{e^x}} = 1 - \frac{e^x}{e^x + 1} = \frac{e^x + 1}{e^x + 1} - \frac{e^x}{e^x + 1} = \frac{1}{e^x + 1}.$$

Тогда

$$1 - \frac{1}{1 + e^{-x}} = \frac{1}{1 + e^x}, \text{ или } 1 - f(x) = f(-x). \quad (4)$$

Будем рассматривать функцию  $f(x)$  только для положительных аргументов, отрицательные их значения можно найти по формуле (4), что позволит ускорить вычисления и уменьшит используемый ресурс ПЛИС.

Э т а п 2. Задаем кусочно-линейную функцию на каждом из интервалов  $(-\infty; x_1)$ ,  $(x_1; x_2)$ , ...  $(x_n; +\infty)$  в виде

$$f(x) = \begin{cases} k_0x + b_0, & x < x_1, \\ k_1x + b_1, & x_1 < x < x_2, \\ \dots & \dots \\ k_nx + b_n, & x_n < x, \end{cases} \quad (5)$$

где  $k$ ,  $b$  — коэффициенты линейного уравнения;  $x$  — аргумент функции активации.

Э т а п 3. Рассчитываем значение  $f(x)$  при ранее полученном значении  $x$ . Из памяти выбираем значения коэффициентов  $k$  и  $b$  для отрицательных аргументов, а затем рассчитываем значение функции (5) по формуле (4).

Таким образом, на основе линейных формул можно найти аппроксимацию функции сигмоидального типа для любого аргумента с заданной точностью.

Алгоритм реализации классической сигмоидальной функции на ПЛИС с использованием языка VHDL.

Ш а г 1. Задаем синаптические веса искусственного нейрона. Каждому нейрону задаем блок оперативной памяти, где хранятся веса. Для задания весов каждого нейрона используются следующие сигналы:

`synaddr` (synapse address) — выбор синапса, вес которого считывается или записывается по его номеру в двоичном коде. Для нейрона из четырех синапсов этот вход двухбитный, из восьми синапсов — трехбитный и так далее;

`synsetw` (synapse set weight) — значение веса, записываемое в синапс;

`synwren` (synapse write enabled) — при наличии единицы на входе нейрон записывает в выбранный синапс значение сигнала `synsetw`, при наличии нуля ничего не происходит.

Ш а г 2. На входы искусственного нейрона подаем значения входного вектора, задаем переменную  $x$ , равную выходу сумматора:  $x = \sum_{i=1}^N w_i a_i$ , где

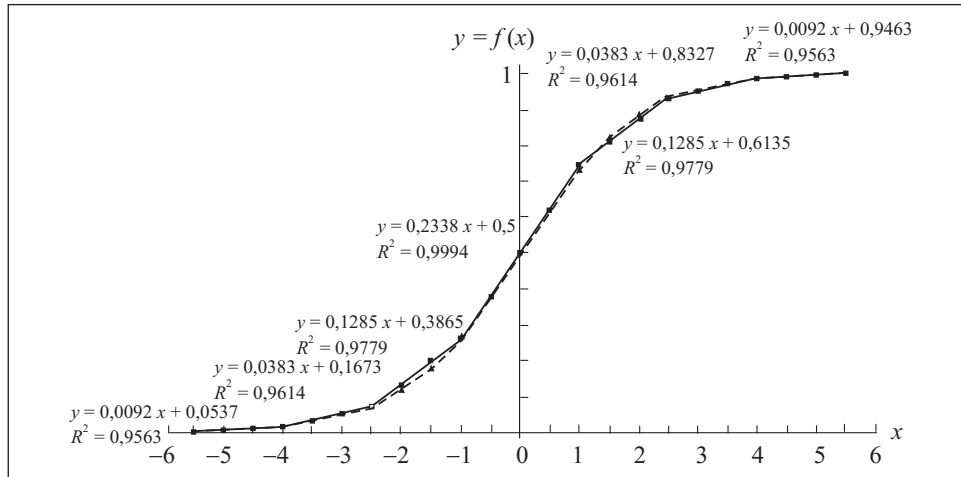


Рис. 3. Кусочно-линейная аппроксимация сигмоидальной функции

$a$  — входы нейрона;  $w$  — синаптические веса нейрона. При расчете используются числа с фиксированной точкой. Каждое число составляет 16 бит, из которых девять приходится на целую часть и семь на дробную. Для арифметических операций над этими числами используется синтезируемая библиотека `fixed_pkg_c.vhdl`, арифметические операции которой хорошо оптимизированы.

Шаг 3. Рассчитываем модуль от аргумента сигмоидальной функции по заданной переменной  $x' = |x|$ .

Шаг 4. Выполняем разбиение на линейные куски сигмоидальную функцию активации и определяем коэффициенты линейных уравнений  $k$  и  $b$  (рис. 3). Определяем коэффициенты линейных уравнений:

$$k, b = \begin{cases} [0,234; 0,500], & 0 \leq x' < 1, \\ [1,129; 0,605], & 1 \leq x' < 2,5, \\ [0,234; 0,500], & 2,5 \leq x' < 4, \\ [0,009, 0,946], & x' \geq 4. \end{cases}$$

Шаг 5. Определяем переменную  $f$ , вычисляя ее значение по формуле  $f = kx' + b$ .

Шаг 6. Если  $x < 0$ , значение локальной переменной изменяется:  $f = 1 - f$ .

Шаг 7. Задаем значение выходного сигнала нейрона равным значению переменной  $f$ .

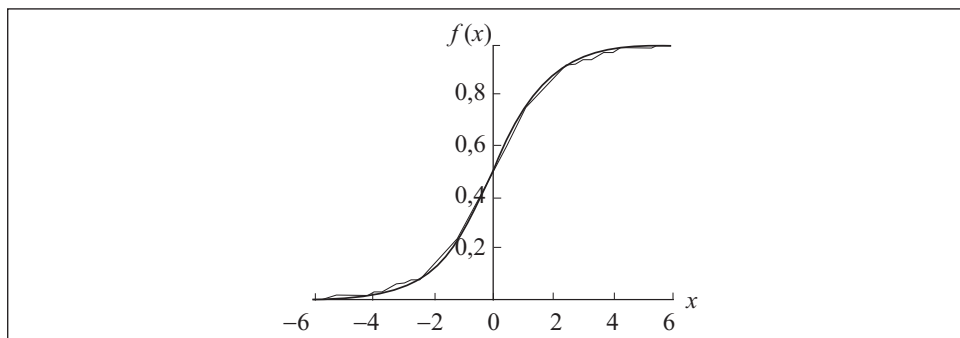


Рис. 4. Сигмоидальная функция, реализованная в ПЛИС

Реализованный по данному алгоритму искусственный нейрон с четырьмя входами и сигмоидальной функцией активации на ПЛИС с использованием 16-разрядных чисел и фиксированной точкой занял объем 672 единицы LUT (Look-Up Table — поисковая таблица) с абсолютной погрешностью  $\pm 0,005$ . Точность реализации сигмоидальной функции представлена на рис. 4. Быстродействие как суммарная задержка комбинационной схемы блока нейрона составила 75,6 нс. При уменьшении или увеличении числа линейных отрезков согласно (5), а также числа входов нейрона или при изменении разрядности чисел будет изменен и использованный ресурс ПЛИС для одного нейрона, а также точность и быстродействие.

Рассмотрим реализацию искусственных нейронных сетей на ПЛИС. Одной из главных особенностей нейросетей является параллельная обработка сигналов. Многослойные нейронные сети представляют собой однородную вычислительную среду. В терминологии нейроинформатики это — универсальные параллельные вычислительные структуры, предназначенные для решения различных классов задач. При аппаратной реализации в ПЛИС каждый нейрон сети является отдельным вычислительным блоком (рис. 5). Нейронная сеть, состоящая из отдельных блоков, представляет собой парал-

Нейронная сеть	Ресурс ПЛИС (LUT)	Значение на выходе		Погрешность	Быстродействие, нс
		Расчетное	Модельное		
1-1-1	1344	0,5626879	0,546875	0,0158129	144,706
1-2-1	2402	0,62343554	0,609375	0,01406054	144,743
2-2-1	2402	0,63427442	0,609375	0,02489942	144,743
2-3-1	3206	0,6954868	0,671875	0,0236118	163,685

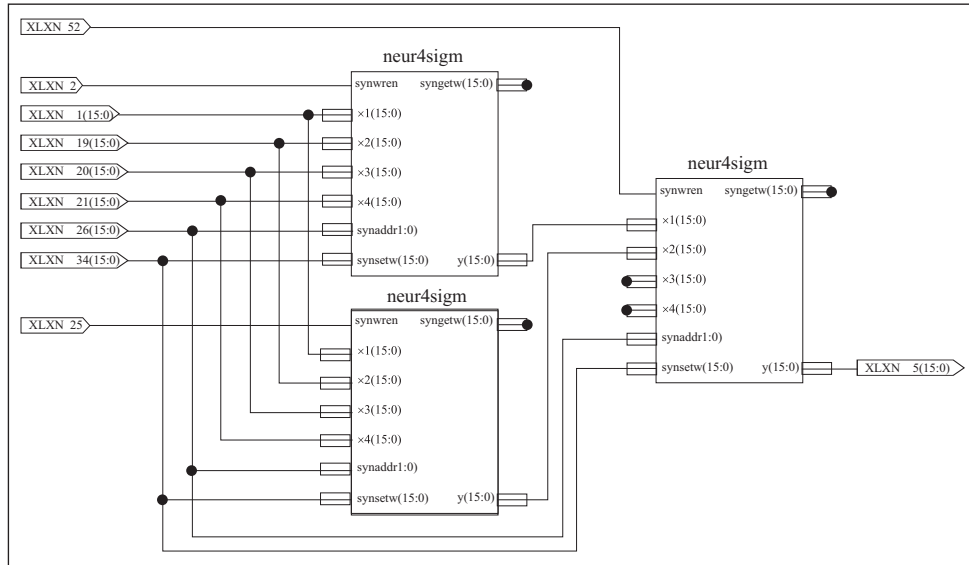


Рис. 5. Нейронная сеть на ПЛИС

тельную вычислительную структуру. Для запуска нейрона используется входной сигнал этого нейрона.

Результаты моделирования нейросети прямого распространения приведены в таблице, где нейронная сеть представлена тремя числами: первое — это число нейронов во входном слое, второе — число нейронов в скрытом слое и третье — число нейронов в выходном слое. Приведены также значения на выходе из нейросети, полученные в результате математического расчета, и в результате моделирования данной нейросети на ПЛИС. Погрешность представляет собой разницу между этими значениями, а быстродействие — суммарную задержку нейронной сети в ПЛИС.

Были также промоделированы нейронные сети Хопфилда и RBF-сети. Аппаратная реализация нейронных сетей Хопфилда на ПЛИС отличается от аппаратной реализации многослойных нейронных сетей прямого распространения введением дополнительных обратных связей и элементов задержки во времени на этих связях, что, в свою очередь, увеличивает используемый ресурс ПЛИС. Алгоритм аппаратной реализации RBF-сетей, в том числе их скрытого слоя радиально-базисных функций Гаусса, выполняется согласно предложенному методу, с той лишь разницей, что функция Гаусса записывается через сигмоидальную функцию [11, 12]. Моделирование выполнено на программном обеспечении Xilinx ISE Design Suite 13.2 и чипе семейства Spartan 3-XC3S200 в ISE Simulator (ISim).



## Выводы

В результате проведенных исследований нейронных сетей при их аппаратной реализации на ПЛИС по предложенному методу достигнута высокая скорость функционирования искусственного нейрона и нейронных сетей с помощью распараллеливания вычислений в самом нейроне и нейросети в целом. Высокая точность и оптимальное использование ресурсов ПЛИС позволят синтезировать на их основе нейросетевые системы управления, функционирующие в одном темпе с процессом управления нелинейными динамическими объектами, которые функционируют в условиях структурной, параметрической и информационной неопределенности.

## СПИСОК ЛИТЕРАТУРЫ

1. Егупов Н.Д. Методы робастного, нейро-нечеткого и адаптивного управления. — М. : Изд-во МГТУ им. Н.Э. Баумана, 2001. — 744 с.
2. Терехов В.А. Нейросетевые системы управления: Учеб. пособие для вузов. — М. : Высшая школа. 2002. — 183 с.
3. Соловьев В. Проектирование цифровых систем на основе ПЛИС. — М. : Радио и связь, 2003. — 376 с.
4. Гильгурт С.Я. Анализ применения реконфигурируемых вычислителей на базе ПЛИС для реализации нейронных сетей // Зб. наук. праць ІПМЕ НАН України «Моделювання та інформаційні технології» — Вип. 37. — Київ: ІПМЕ НАН України, 2006. — С. 168—174.
5. Сергиенко А.М. VHDL для проектирования вычислительных устройств — Киев: ЧП «Корнейчук», ООО «ТИД «ДС», 2003. — 208 с.
6. Сергиенко А.М., Симоненко В.П. Отображение периодических алгоритмов в программируемые логические интегральные схемы // Электрон. моделирование. — 2007. — 29. — № 2. — С. 49—61.
7. Kravets P.I., Lukina T.I., Zherebko V.A., Shimkovich V.N. Methods of Hardware and Software Realization of Adaptive Neural Network PID Controller on FPGA-Chip. // J. of Automation and Information Sciences. — 2011, Vol. 43, Issue 4. — P. 70—77.
8. Кравець П.І., Лукіна Т.Й., Шимкович В.М., Ткач І.І. Розробка та дослідження технології оцінювання показників нейромережових моделей МІМО-об'єктів управління // Вісн. НТУУ «КПІ». Інформатика, управління та обчислювальна техніка: Зб. наук. праць. — 2012. — № 57. — С. 144—150.
9. Кравець П.І., Шимкович В.Н. Метод оптимізації весових коефіцієнтів нейронних мереж з допомогою генетичного алгоритму при реалізації на програмуємих логічних інтегральних схемах //Електрон. моделювання. — 2013. — 35, № 3. — С. 65—75.
10. Кравець П.І., Шимкович В.М., Омельченко П. Нейромережові компоненти систем керування динамічними об'єктами з їх апаратно-програмною реалізацією на FPGA / Вісн. НТУУ «КПІ». Інформатика, управління та обчислювальна техніка: Зб. наук. праць. — 2013. — № 59. — С. 78—85.
11. Кравець П.І., Шимкович В.М., Зубенко Г.А. Технологія апаратно-програмної реалізації штучного нейрона та штучних нейронних мереж засобами FPGA // Там же. — 2012. — № 55. — С. 174—180.

12. Кравець П.І., Шимкович В.М., Зубенко Г.А. Моделі штучних нейронних мереж при їх апаратно-програмній реалізації на FPGA / Сб. тр. XIV міжнародної науч. конф. ім. Т.А.Таран «Інтелектуальний аналіз інформації ІАІ-2014», Київ, 14—16 мая 2014 г. — Київ : Просвіта, 2014. — С. 127—131.

*P.I. Kravets, V.N. Shimkovich, D.A.Ferens*

#### METHOD AND ALGORITHMS OF IMPLEMENTATION ON PLIS THE ACTIVATION FUNCTION FOR ARTIFICIAL NEURON CHAINS

A method and algorithms of the hardware-software implementation of neuron networks on PLIS of FPGA type using VHDL language is proposed in the work. Examples of the hardware-software implementation by PLIS means of artificial neurons and neuron networks with sigmoid activation functions are presented; it is shown that the considerable velocity of data processing is provided and the use of PLIS resource is optimized owing to the developed method and algorithm.

*Key words: neuron networks, activation functions, PLIS, VHDL language.*

#### REFERENCES

1. Yegupov, N.D. (2001), *Metody robastnogo, neuro-nechetkogo i adaptivnogo upravleniya* [Methods of robust, neuro-illegible and adaptive control], Izdatelstvo MGTU im. N.E. Bauman, Moscow, Russia.
2. Terekhov, V.A. (2002), *Neirosetevye sistemy upravleniya: Ucheb.Posobie dlya vuzov* [Neuronetwork systems of control: Manual for Higher Educ. Inst.], Vysshaya shkola, Moscow, Russia.
3. Soloviev, V. (2003), *Proektirovanie tsifrovyykh system na osnove PLIS* [Design of digital systems on PLIS basis], Radio i svyaz, Moscow, Russia.
4. Gilgurt, S.Ya. (2006), “Analysis of the use of reconfigurable computers based on PLIS for implementation of neuron networks”, *Modelyuvannya ta informatsiini tekhnologii. Zb. Nauk.Pr. IPME NAN Ukrainy*, Iss. 37, pp. 168-174.
5. Sergienko, A.M. (2003), *VHDL dlya proektirovaniya vychislitelnykh ustroystv* [VHDL for design of computing devices], ChP “Korneichuk”, OOO “TID “DS”, Kiev, Ukraine.
6. Sergienko, A.M. and Simonenko, V.P. (2007), “Mapping of periodic algorithms to programmed logical integral circuits”, *Elektronnoe modelirovanie*, Vol. 29, no. 2, pp. 49-61.
7. Kravets, P.I., Lukina, T.I., Zherebko, V.A. and Shimkovich, V.N. (2011), “Methods of Hardware and Software Realization of Adaptive Neural Network PID Controller on FPGA-Chip”, *Journal of Automation and Information Sciences*, Vol. 43, Iss. 4, pp. 70-77.
8. Kravets, P.I., Lukina, T.I., Shimkovich, V.N. and Tkach, I.I. (2012), “Development and study of the process of estimation of indices of the control MOMO-objects models”, *Visnyk NTUU “KPI”, Informatyka, upravlinnya ta computer engineering: Zb. nauk. Prats*, no. 57, pp. 144-150.
9. Kravets, P.I. and Shimkovich, V.N. (2013), “Method of optimization of weight coefficients of neuron networks by means of genetic algorithm under implementation on programmed logical integral circuits”, *Elektronnoe modelirovanie*, Vol. 35, no. 3, pp. 65-75.
10. Kravets, P.I., Shimkovich, V.N. and Omelchenko, P. (2013), “Neuronetwork components of the systems of control of dynamic objects and their hardware-software implementation on FPGA”, *Visnyk NTUU “KPI”, Informatyka, upravlinnya ta computer engineering: Zb. nauk. Prats*, no. 59, pp. 78-85.

11. Kravets, P.I., Shimkovich, V.N. and Zubenko, G.A. (2012), “Technology of hardware-software implementation of artificial neuron and artificial neuron networks by FPGA method”, *Visnyk NTUU “KPI”, Informatyka, upravlinnya ta computer engineering*, no. 55, pp. 174-180.
12. Kravets, P.I., Shimkovich, V.N. and Zubenko, G.A. (2014), “Models of artificial neuron networks under their hardware-software implementation on FPGA”, *Intellectualnyi analiz informatsii LAI-2014. Sbornik trudov XIV mezhdunarodnoi nauchnoi konferentsii im. T.A. Taran* [Intellectual analysis of information. Collection of proceedings of T.A. Taran XIV International Conference], Kiev, May 14-16, 2014, pp. 127-131.

Поступила 21.05.15;  
после доработки 24.06.15

*КРАВЕЦ Петр Иванович, канд. техн. наук, доцент кафедры автоматики и управления в технических системах Национального технического университета Украины «КПИ». В 1972 г. окончил Национальный технический университет Украины «Киевский политехнический институт». Область научных исследований — автоматизированные системы управления технологическими процессами.*

*ШИМКОВИЧ Владимир Николаевич, аспирант, ассистент кафедры автоматики и управления в технических системах Национального технического университета Украины «КПИ». В 2010 г. окончил Национальный технический университет Украины «Киевский политехнический институт». Область научных исследований — автоматизированные системы управления технологическими процессами.*

*ФЕРЕНС Дмитрий Андреевич, студент Национального технического университета Украины «КПИ», кафедра автоматики и управления в технических системах.*

