

## **К РЕАЛИЗАЦИИ МОДИФИЦИРОВАННОЙ ВЕРОЯТНОСТНОЙ НЕЙРОННОЙ СЕТИ АДАПТИВНОГО РЕЗОНАНСА**

**Д.В. Постарнак**

аспирант, e-mail: dmi1852@yandex.ru

**В.А. Шапцев**

д.т.н., профессор кафедры информационных систем, e-mail: VAShaptsev@yandex.ru

Тюменский государственный университет

**Аннотация.** Предлагается вариант эффективного распределения производительности между вычислительными компонентами при выполнении распознавания образов методом выбора направления в модифицированной вероятностной нейронной сети адаптивного резонанса.

**Ключевые слова:** нейрон, теория адаптивного резонанса, слой распознавания, синаптический вес, процессор, параллельные вычисления.

### **Введение**

В реализации метода распознавания образов, описанного в [1], производительность и надёжность системы процесса можно повысить применением схемы параллельных вычислений. Ниже представлены модель искусственной нейронной сети, отличающаяся включением выбора направления поиска, и вариант распределения производительности при параллельных вычислениях в процессе распознавания.

### **1. О модели нейронной сети**

Рассматривается модель АРТ (от англ. Adaptive Resonance Theory) нейронной сети, разработанная Стивеном Гроссбергом [2]. Основные черты, отличающие её от остальных: самостоятельное обучение и повышение качества распознавания за счёт наполнения слоя распознающих нейронов [3]. Важны следующие особенности этой модели:

- конкурентный механизм выбора сработавшего нейрона;
- определение степени различия между идеальным образом, закреплённым за текущим распознающим нейроном, и текущим входным образом. На основании этого выносится решение о том, считать состояние текущего нейрона, участвующего в распознавания, результатом классификации или нет;

- если ни один из распознающих нейронов не обеспечивает необходимого уровня соответствия входному образу, то происходит «адаптивный резонанс» - включение в слой распознавания нового нейрона, несущего идеальный образ, полностью соответствующий текущему входному.

В [1] предложена вероятностная модель АРТ, названная там PART (Probabilistic ART). Одной из особенностей этой модификации является использование нескольких связанных структур в качестве слоя распознавания. Это позволяет учесть нюансы входных образов, упускаемые другими типами моделей, и использовать эти нюансы для определения вероятностного показателя успешности пути обхода дерева [1]. Теперь классификация ведётся не по одному нейрону, а по набору сработавших нейронов (порядок расположения которых неважен). Процесс сопровождается набором результатов классификации, каждый (набор) из которых представлен порядковыми номерами нейронов распознавания:

$$B = \langle C, D \rangle, \quad (1)$$

где  $C$  — набор порядковых номеров распознавших нейронов, которые формируют успешную текущую классификацию  $B$ ;  $D$  — число успешных классификаций с данным результатом. Вводим множество

$$A = \{B_1, B_2, \dots, B_n\}, \quad n \geq 0, \quad (2)$$

где  $B_i$  —  $i$ -й результат классификации;  $n$  — число результатов классификации в системе распознавания.

Безуспешная классификация определяется как пустое множество *сработавших* нейронов:  $C = \emptyset$ . В этом случае, при использовании *скалярного вида образа* (СВО) [1], нецелесообразно переходить к следующему узлу ветвления. Можно вернуться к предыдущему нейрону и по другим СВО попробовать достроить *полосу классификации* с другим окончанием. Если от предыдущего нейрона ни один из ближайших (допустимых) СВО не даёт результата, то спускаемся к ещё более раннему распознающему нейрону и проделываем то же самое. Если вплоть до самого корня классификация не удалась, то возвращаемся к первоначальному виду невыполненной классификации и повторяем процесс постепенного перехода к предыдущим нейронам без учёта СВО. Если и это не помогло, то делается вывод о том, что классификация от текущего корня не состоялась.

## 2. Адаптация к параллельным вычислениям

Алгоритм параллельных вычислений в рамках PART по своей сути не меняется. Разница состоит в том, что теперь процесс ведётся не только от корня к листьям, но и в обратном направлении. Совокупность компонентов, закреплённая за распознающим нейроном, меняет вид:

$$Q = \langle G, I, E_1, E_2, J, M \rangle, \quad (3)$$

где  $G$  — набор порядковых номеров распознающих нейронов, связанных с текущим. Он используется при движении от корня к листьям;  $I$  — вектор, содержащий набор порядковых номеров распознающих нейронов, связанных с текущим. Используется при движении от листа к корню;  $E_1$  — СВО при движении от корня к листьям;  $E_2$  — СВО при движении от листа к корню;  $M$  — маркер участия нейрона в текущей классификации.

Вектор  $J$  определяет набор входов, весовые коэффициенты связи нейрона слоя распознавания со входом и элементы образа, на которые настроен нейрон (идеального образа) по отношению к соответствующим входам:

$$J = (l_1, b_1, t_1; l_2, b_2, t_2; \dots; l_n, b_n, t_n), n \geq 0, \quad (4)$$

где  $l_i$  ( $i = 1, 2, \dots, n$ ) — порядковый номер входа;  $b_i$  ( $i = 1, 2, \dots, n$ ) — весовой коэффициент связи нейрона слоя распознавания со входом;  $t_i$  ( $i = 1, 2, \dots, n$ ) — элемент образа, на который настроен нейрон;  $n$  — число входов. Наборы могут отличаться по размеру при условии, что они не перекрываются и полностью охватываются конкурирующими нейронами.

Теперь кроме набора порядковых номеров нейронов со связью вверх, то есть от корня к листу, появляется такой же набор со связью вниз. Скалярный вид образа (СВО) также представлен как для верхнего хода, так и для нижнего. Вычисляется СВО аналогично как вверх, так и вниз. То есть суть учёта нюансов не меняется. Есть ещё «маркер», указывающий на то, был ли использован конкретный нейрон в текущем процессе распознавания, т.е. маркер может быть равен единице в одном случае, а в следующий раз — нулю.

Система распознавания PART с применением параллельных вычислений предполагает использование дополнительных структур данных. Листья хранятся отдельно

$$O = (P_1, P_2, \dots, P_n), n \geq 0, \quad (5)$$

где  $P_i$  — набор порядковых номеров нейронов распознавания, используемых в качестве листьев;  $i$  — номер корня, которому соответствует набор.

Корни (возможные начала классификаций) должны быть также отмечены во внешнем векторе:

$$N = (m_1, m_2, \dots, m_n), n \geq 0, \quad (6)$$

где  $m_i$  — порядковый номер нейрона распознавания, используемого в качестве корня;  $i$  — номер корня в общем наборе.

Хранение данных о корнях и листьях необходимо для релевантного распределения производительности во избежание ситуации, когда выбирается несколько нейронов, идущих из одного корня и, как следствие, идёт проверка одного и того же образа несколько раз подряд, сводя на нет смысл параллельной работы.

Всякий раз, когда процесс классификации доходит до листа, у которого имеется только обратный набор ссылок, а прямой набор ссылок пуст, этот последний элемент (лист) заносится в вектор листьев (получается, что при условии его отсутствия в нём).

### 3. Работа модели при параллельных вычислениях

Параллелизм работы модели заключается в следующем. Выбирается любой корень из (6). По виду вектора (6) можно определить, какие листья соответствуют выбранному корню (порядок следования наборов листьев  $P$  и корней  $m$  совпадает). Запускается классификация от выбранного корня к листьям на первом процессоре CPU1 (рис. 1). По крайним листьям с двух сторон запускается классификация на двух процессорах (два процессора идут по листьям друг к другу) (верхняя часть рис. 1). Если один процессор CPU3 «входит» в дерево, то второй CPU2 не останавливается. Его точка входа может быть ближе к краю дерева. Нормально, что оба процессора CPU2 и CPU3 включились в работу. Главное, чтобы они начали вычисления с разных листьев.

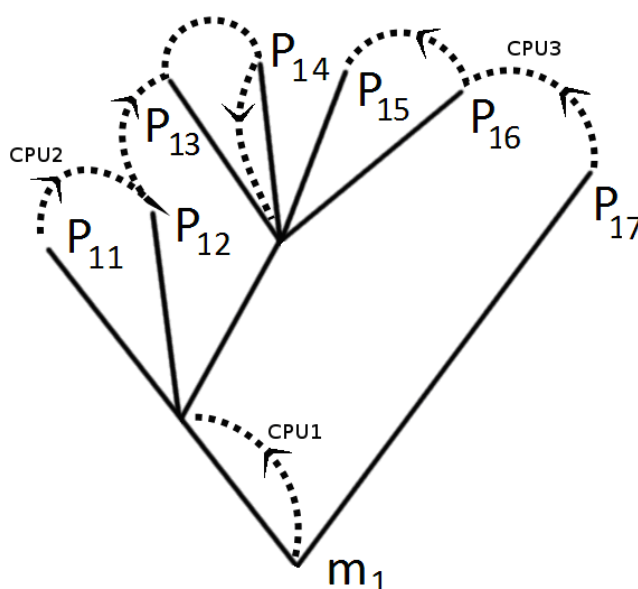


Рис. 1. Вход в дерево

Назовём *ведомыми процессорами* те, которые обходят листья, как описано выше, то есть — CPU2 ( $P_{11}, P_{12}, P_{13}, P_{14}$ ) и CPU3 ( $P_{17}, P_{16}, P_{15}$ ). Они при погружении должны использовать информацию о СВО.

Если *решающая полоса* (последовательность переходов от нейрона к нейрону, формирующая результат классификации) *ведущего процессора* (того, который ведёт обработку от корня к листьям, в нашем случае CPU1) не соединилась ни с одной из *решающих полос ведомых процессоров* (смотрим по маркеру), то в самом худшем варианте у нас есть три (по общему числу процессоров) завершённых классификации (которые были пройдены по правилам АРТ-концепции) (рис. 2).

В этом случае берём ту полосу ( $m_1 - n_1 - P_{11}$  или  $m_1 - n_1 - n_2 - P_{14}$  или  $P_{17} - P_{16} - P_{15}$ ), у которой самое большое число срабатываний. Число

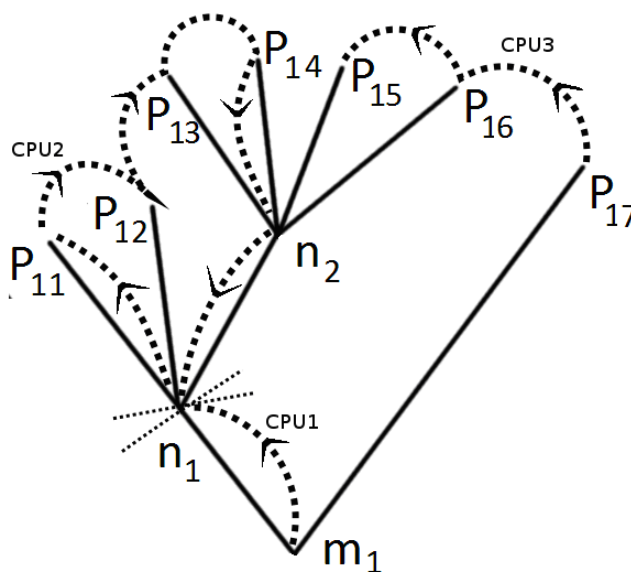


Рис. 2. Расхождение параллельных процессов

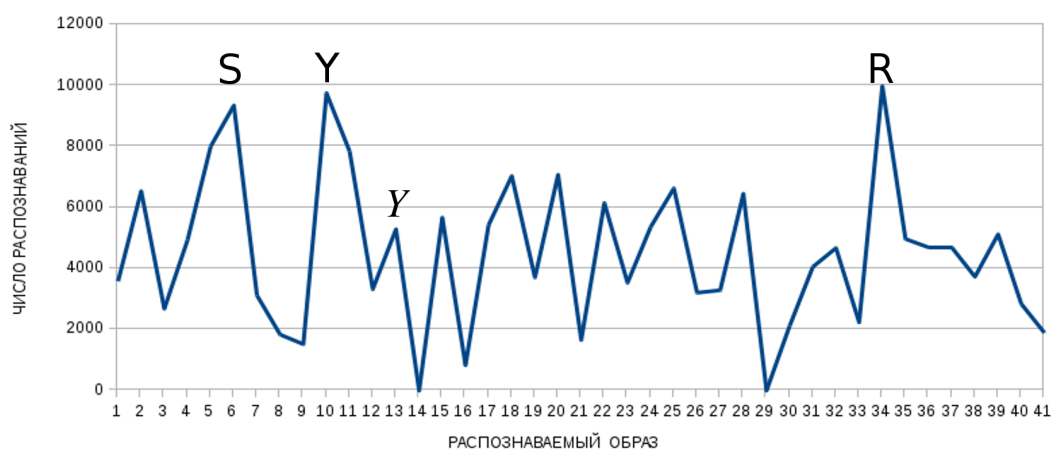


Рис. 3. Число распознаваний образов

срабатываний (в рамках модели  $A$ ) (2) увеличивается только тогда, когда есть все сработавшие полосы с одинаковыми показателями по срабатыванию (рост только у одной из полос  $B$ ) (1). Следовательно, на пик по числу срабатываний выходят те образы, которые между собой не соперничали (или соперничали менее интенсивно, чем остальные). Это показано на рис. 3: два варианта написания символа «Y» могли конкурировать до того момента, когда количество символов «Y», написанных курсивом, заметно убавилось.

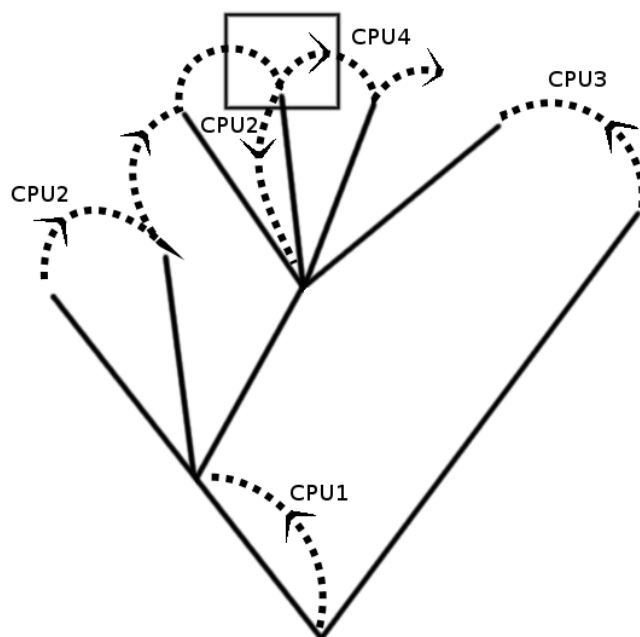


Рис. 4. Продолжение прохода по листьям

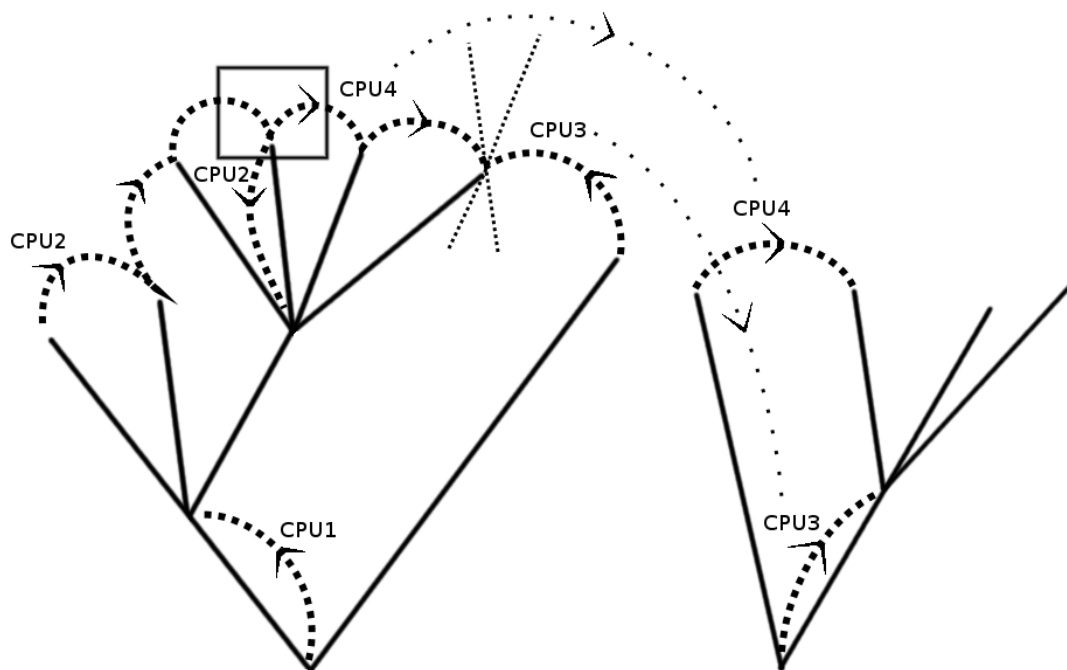


Рис. 5. Параллельное следование от корня

Если один из ведомых процессоров «входит» в дерево, то можно продолжить проход по листьям в том же направлении другими процессорами (рис. 4).

При встрече идущих по листьям процессоров некоторые из них могут освободиться. Из освободившихся процессоров можно сформировать новую группу «ведущий — два ведомых» или «ведущий — ведомый», или вообще один ведущий, лишь бы процессоры не простаивали (рис. 5).

## Заключение

Выше представлен вариант эффективного (в смысле ускорения вычислений) распределения производительности между процессорами при выполнении распознавания образов методом выбора направления. Приведённая методика сохраняет эффективность системы при частичном выходе из строя вычислительных компонентов, то есть система может менять число используемых процессоров (по мере их доступности) во время непрерывной работы. Производительность системы должна возрасти в число процессоров за минусом времени, затраченного на управление распараллеливанием.

## ЛИТЕРАТУРА

1. Постарнак Д.В. Метод выбора направления для ускорения классификации с помощью вероятностной нейронной сети адаптивного резонанса // Математическое и информационное моделирование: сб. науч. тр. 2013. № 13. С. 218–229.
2. Grossberg S. Competitive Learning: From Interactive Activation to Adaptive Resonance // Cognitive science. 1987. № 11. P. 23–63.
3. Постарнак Д.В. Критический анализ моделей нейронных сетей // Вестник ТюмГУ. 2012. № 4. С. 162–167.

## TOWARDS IMPLEMENTING THE MODIFIED PROBABILISTIC ADAPTIVE RESONANCE NEURAL NETWORK

**D.V. Postarnak**

Postgraduate Student, e-mail: dmi1852@yandex.ru

**V.A. Shaptsev**

Professor, Doctor of Engineering, e-mail: VAShaptsev@yandex.ru

Tumen State University

**Abstract.** We propose a method for an efficient allocation of performance between computational components when carrying out the pattern recognition by choosing direction in the modified probabilistic adaptive resonance neural network.

**Keywords:** neuron, adaptive resonance theory, recognition layer, synaptic weight, processor, parallel computing.