

МЕТОДИКА СТРУКТУРНОГО ПРОЕКТИРОВАНИЯ ЭЛЕКТРОННЫХ НЕЙРОСЕТЕЙ

Алюшин А.В., Алюшин М.В., Алюшин С.А.

Московский инженерно-физический институт (Государственный университет)
115409, Каширское шоссе, 31

Реферат. Предложена методика структурного проектирования электронных нейросетей для решения определенного класса задач, позволяющая определить ее основные параметры (число слоев, количество нейронов в каждом слое, тип связей между слоями), уточнить требования к конструкции нейрона, снизить требования к интегральной технологии изготовления.

Введение

При решении практических задач обработки сигналов разработчики электронных нейроподобных систем вынуждены использовать сети заведомо больших размеров, перебирать различные нейросетевые парадигмы с целью получения приемлемых результатов, так как структура сети априори неизвестна. Структурная оптимизация сети особенно важна при разработке интегральных систем обработки сигналов.

В данной работе предложена методика структурного проектирования электронных нейросетей для оптимизации размеров сети, определения параметров ее элементов.

Искусственная нейросеть с обобщенной структурой

Данный тип нейросети представляет собой M входных шин, N выходных шин, «линейку» из N нейронов, выход каждого нейрона сети соединен с соответствующей выходной шиной сети, со входом каждого нейрона сети через соответствующий синаптический элемент W_{ij} . Кроме того, вход каждого нейрона соединен с каждой входной шиной через соответствующий синаптический элемент с весом V_{ij} (см. рис.1).

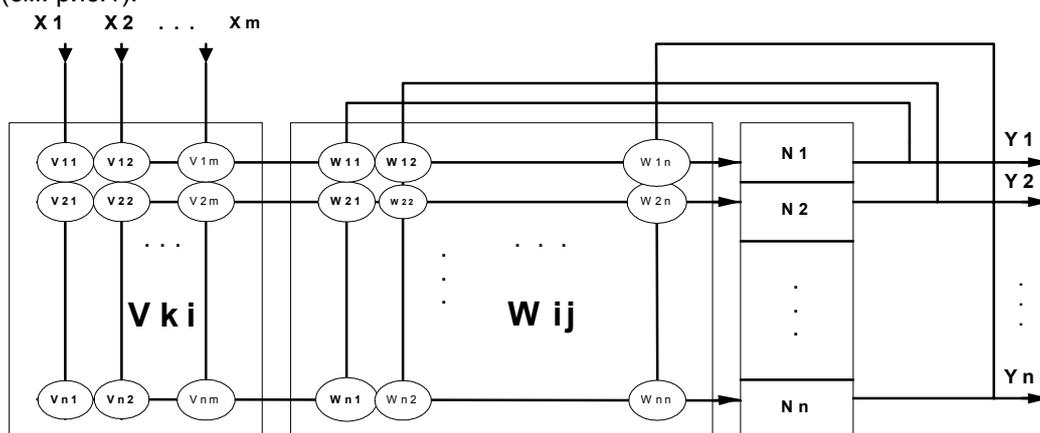


Рис. 1. Нейросеть с обобщенной структурой

Данная нейросеть является универсальной, ее структура не настроена на решение определенного класса задач. Прямая реализация подобной сети в электронном виде при $N=10^3 - 10^6$ и небольших M потребует реализации соответственно $10^6 - 10^{12}$ синаптических элементов, что предъявляет повышенные требования к технологии изготовления.

Целью структурного проектирования является оптимизация структуры сети для эффективного решения определенного класса задач. Последовательность действий при структурном проектировании нейросети:

- создать искусственную нейросеть с обобщенной структурой достаточного размера (например, в виде программы);
- обучить нейросеть на ограниченном наборе примеров, достаточно полно характеризующих выбранный класс задач;
- осуществить кластеризацию элементов полученных после начального обучения матриц V_{kj} и W_{ij} (группировка весовых коэффициентов с большими значениями в один или несколько кластеров с возможно меньшей суммарной площадью этих кластеров за счет индексных преобразований матриц);
- осуществить нелинейное преобразование элементов матриц V_{kj} и W_{ij} , например, с помощью пороговой функции (обнуление весовых коэффициентов с небольшими значениями);

- аппроксимировать полученные кластеры непересекающимися квадратными областями;
- определить количество (число слоев в сети), размеры (количество нейронов в данном слое сети) и положение (тип связей между слоями и в пределах одного слоя) квадратных областей, содержащих ненулевые весовые коэффициенты матрицы;
- определить равномерность распределения значений весовых коэффициентов в пределах одного кластера (информация о требуемой разрядности весовых коэффициентов),
- синтезировать электронную нейросеть с оптимизированной структурой.

Влияние ошибки нелинейного преобразования матриц на точность решения задач выбранного класса в определенной степени компенсируется обучением сети с оптимизированной структурой на расширенном наборе примеров после ее разработки и изготовления.

Форма представления сигналов в электронной нейросети (аналоговая, цифровая, аналоговая и цифровая) определяется выбранным схемотехническим базисом.

В качестве примера на рис.2 представлен вариант матрицы W_{ij} , соответствующий двухслойной сети прямого распространения с количеством нейронов в каждом слое соответственно $N_1 = 3$, $N_2=3$.

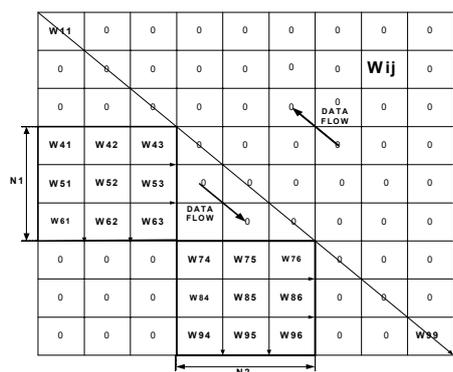


Рис. 2. Сеть прямого распространения

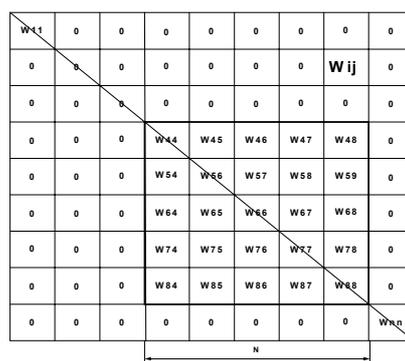


Рис. 3. Сеть Хопфилда

Заметим, что при данной нумерации элементов матрицы кластер, расположенный снизу от диагонали матрицы, отвечает за слой нейронов с прямым распространением сигналов в сети. Кластер, расположенный над диагональю матрицы, отвечает за слой нейронов с обратным распространением сигналов в сети. На рис. 3 приведен пример матрицы, соответствующий одному слою нейронов с полными перекрестными связями при $N = 5$ (сеть Hopfilda [1]). На рис. 4 представлен вариант распределения весовых коэффициентов для систолической сети с размерами соответственно $A \times B=16$.

Очевидно, что этап кластеризации весовых коэффициентов может привести к различным топологическим вариантам, сильно зависит от начальной инициализации матриц V_{kj} и W_{ij} . Определенную помощь в проведении кластеризации могут оказать сведения о структуре биологической нейросети – прототипа разрабатываемой системы. Более того, разработка определенным образом синтезированных звуковых или видео тестов и проведение экспериментальных исследований в случае создания системы распознавания голосовых или зрительных образов позволит получить дополнительную информацию о структуре нейросети живых биологических объектов [2]. В случае, если структура связей биологической сети-аналога известна достаточно точно, этап структурного проектирования электронной нейросети также целесообразен, так как функции искусственного нейрона в любом случае отличаются от биологического. Вариация некоторых параметров исходной нейросети с обобщенной структурой (тип функции нелинейного преобразования, способ генерации выходного сигнала нейрона, задержка распространения сигнала для каждой синаптической связи) также оказывает существенное влияние на процесс кластеризации. В этом случае, можно выполнить следующие шаги на начальном этапе структурного проектирования - инициализировать матрицы начальными значениями, форма распределения которых может быть отнесена к одному из стандартных видов или их суперпозиции (см. например, рис.1-3). После обучения на небольшом наборе примеров проанализировать тенденцию изменения распределения весовых коэффициентов (аналог генетического алгоритма) и определить возможную устойчивую систему путем аппроксимацией квадратами и т.д.

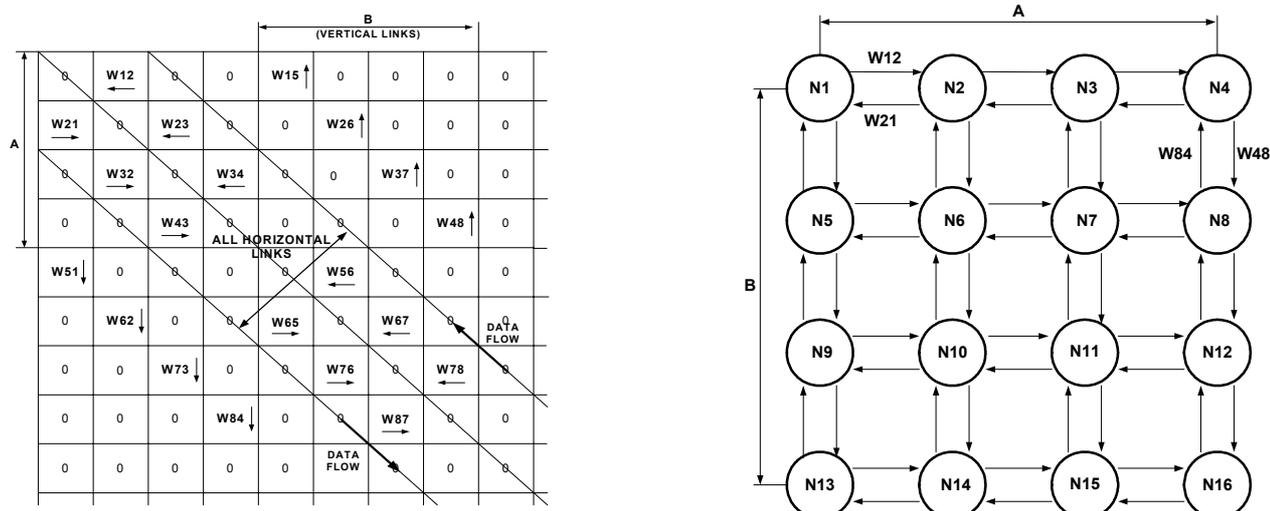


Рис. 4. Систолическая сеть

Выводы

Таким образом, предложенная методика структурного проектирования электронных нейросетей позволяет определить необходимое для решения выбранного класса задач количество слоев сети, количество нейронов в каждом слое, типы связей между слоями, входными и выходными шинами, оценить точность представления весовых коэффициентов, степень влияния параметров нейрона.

Полученная информация о структуре сети особенно важна при разработке микроэлектронных систем обработки сигналов.

Литература

1. Hopfield J.J. Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of sciences, USA, Vol. 79, p.2554-2558, 1982.
2. Алюшин С.А. Методика экспериментального определения структурных параметров нейросети для распознавания речи. //Молодежь и наука: Сб. научных трудов. - М.:МИФИ, 2001, с.84-85.