

Исследование динамических нейронных сетей в интересах проектирования вычислительных сетей заданной топологии (часть 1). Постановка задачи и подход

А.Г. Волков,
к.т.н., доц., volkov-ag@mail.ru,
Финансовый ун-т, г. Москва,
И.Н. Иванова,
вед. инж.-прогр., inivanovaw@mail.ru,
ИПУ РАН, г. Москва,
А.В. Радионов,
ВА ВВСН., г. Балашиха

Исследуются вопросы применения нейросетевых технологий для быстрого проектирования вычислительных сетей. Предложен подход к проектированию вычислительных сетей наземного комплекса управления космическим аппаратом заданной топологии на основе динамических нейронных сетей. Разработанный пошаговый алгоритм позволяет провести нейросетевую интерпретацию задачи проектирования структуры сети заданной конфигурации.

В первой части доклада представлены постановка задачи, общий подход к решению и базовая модель.

The report describes the possibility of using neural network technologies for the rapid design of computer networks. An approach to the design of computer networks for a ground control system of a spacecraft with defined topology based on dynamic neural networks is described. The developed step-by-step algorithm allows for a neural network interpretation of the problem of designing a network structure of the defined configuration.

The first part of the report presents the statement of the problem and common approach to the solution and the basic model.

Введение

Проблемам предпроектных исследований в аэрокосмической области посвящено множество публикаций [1-3]. Многие проблемы принятия решений при разработке и эксплуатации космической техники могут быть представлены в виде дискретных и комбинаторных оптимизационных задач. Одной из таких задач является задача проектирования структуры вычислительной сети наземного комплекса управления (НКУ) с учётом временных и финансовых ограничений. В связи с этим большое практическое значение приобретает развитие способов получения эффективного решения группы таких задач. Анализ развития теории дискретных и комбинаторных оптимизационных задач и практика их решения показывают, что такая задача может быть решена методом полного перебора, который, однако, характеризуется значительной трудоёмкостью [4,5].

Например, в задачах программирования с булевыми переменными, в которых размерность вектора параметров x равна n , а элементы вектора x могут принимать значения $\{0, 1\}$, мощность Δ множества альтернатив будет равна 2^n . Число шагов вычислительного процесса при реализации алгоритма полного перебора и соответственно эффективность данного алгоритма также определяется числом 2^n . Таким образом, в задачах принятия комбинаторных решений применительно к сложным оптимизационным задачам, для которых характерна большая размерность исходных данных, возможности полного перебора весьма ограничены и возникает задача доказательства существования и последующего синтеза алгоритма существенно более экономичного, чем перебор.

Существующие методы, в зависимости от требуемого качества получения результата, можно разделить на два класса [4,6]:

- методы, которые всегда приводят к нахождению оптимального решения, но требуют для этого недопустимо большого числа операций;
- методы, которые не всегда приводят к нахождению оптимального решения, но требуют для получения этого решения приемлемого числа операций.

Ко второму классу можно отнести нейросетевые методы, использующие прием, заключающийся в отказе от поиска оптимального решения и в нахождении рационального решения за требуемое время.

Рассмотрим порядок нейросетевой интерпретации задачи топологического проектирования сети. Этап топологического проектирования предусматривает [5] решение следующих задач:

- определение структуры сети;
- размещение структурных компонентов в пространстве определенном техническим заданием (ТЗ) для получения базовых вариантов топологии системы;
- улучшение, совершенствование базовых топологий для получения систем с определенными в ТЗ функциональными характеристиками.

Объектом моделирования выберем разветвленную сеть, состоящую из концентраторов и подключенных к ним хостов, которые в свою очередь могут являться маршрутизаторами для подсетей второго уровня. При этом каждый хост может быть подключен к любому маршрутизатору. Связи между хостами и концентраторами будем рассматривать как двунаправленные, что соответствует дуплексным каналам связи.

Основные виды структур могут быть разбиты на 4 класса [5]: древовидные; ячеистые; сильносвязанные; смешанные.

Анализ достоинств и недостатков перечисленных структур позволяет сделать вывод о выборе простейшей древовидной структуры в качестве базовой, с последующим наращиванием сети и введением дополнительных каналов, например между концентраторами с целью обеспечения требуемого согласно ТЗ уровня разветвленности.

Для оптимизации структуры древовидных сетей могут использоваться методы исследования операций, которые позволяют найти оптимальное решение, однако с ростом размерности задачи (количество хостов и коммутаторов превышает 100) время нахождения решения растет полиномиально и может достигать значений превышающих время проектирования, установленное ТЗ.

Поэтому на практике часто используются эвристические алгоритмы, дающие рациональные решения за конечное число шагов даже для сетей большой размерности (более 100 элементов). Среди таких алгоритмов можно выделить алгоритмы Прима, Ежи-Вильямса и др. [5]. Однако при дальнейшем увеличении числа элементов сети до 1000, качество решения оптимизации структуры сети может не удовлетворять требованиям технического задания.

Постановка задачи

Одним из возможных путей решения задачи может быть нейросетевая интерпретация точных методов с использованием математического аппарата динамических сетей Хопфилда, использующих понятие энергии сети [7-9]. Энергетическую функцию сети можно представить в виде:

$$E = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^M \sum_{\mu=1}^N \sum_{\nu=1}^M W_{ij\mu\nu} Y_{ij} Y_{\mu\nu} + \sum_{i=1}^N \sum_{j=1}^M Y_{ij} I_{ij} \quad (1)$$

Задачу определения структуры сети древовидной конфигурации сформулируем следующим образом:

Дано:

$A = \{a_i\}$ – где $i = 1 \dots N$ множество узлов проектируемой сети, среди которых узел с номером $i = 1$ соответствует коммутатору, а каждый из узлов с номером $i \geq 2$ характеризуется требуемой интенсивностью λ_i потока сообщений.

$X = \{x_{ij}\}$ – матрица связей, которыми могут быть соединены любая пара узлов a_i, a_j .

$\mathcal{R} = \{m_{ij}\}$ – множество взвешенных длин связей, которые могут быть представлены матрицей $N \times M$. Вес связи определяется затратами на установку коммуникационного оборудования, протяженностью линии связи и т.п.

$B = \{b_{ij}\}$ - множество допустимых пропускных способностей каждой связи.

$P = \{p_{ij}\}$ – допустимая суммарная пропускная способность.

Ограничения:

условие древовидности, состоящее в том, что из каждого узла (кроме $i=1$) выходит только одна связь

$$\sum_{j=1}^M x_{ij} = 1, \quad i \neq 1, \quad (2)$$

$$\sum_{i=1}^N \sum_{j=1}^M x_{ij} = N. \quad (3)$$

ограничение на пропускную способность связей, причем коммутаторы в сети связываются последовательно, а дополнительные связи, повышающие надежность иерархической сети не рассматриваются

$$\sum_{i=1}^N \sum_{j=1}^M B_{ij} X_{ij} = \sum_{j=1}^M P_{j+1}. \quad (4)$$

целевая функция

$$Q = \sum_{i=1}^N \sum_{j=1}^M m_{ij} x_{ij}, \quad (5)$$

где $x_{ij} = 1$, если $m_{ij} \in S_j$ и $x_{ij} = 0$, если $m_{ij} \notin S_j$, где S – множество коммутаторов, а $x_{ij} \in \{0, 1\}$.

Требуется:

Определить конфигурацию сети, удовлетворяющую ограничениям (2-5), такую что:

$$Q^* = \arg \min_{x_{ij}} Q \quad (6)$$

Такая постановка задачи в целом соответствует постановке известной задачи разбиения [4,6].

Решение задачи

Для нейросетевой интерпретации постановки задачи будем рассматривать матрицу бинарных нейронов $[Y_{ij}]$ размерности $N \times M$, возбужденное состояние нейрона в которой соответствует тому факту, что i -ый хост подключен к j -ому коммутатору, т.е. каждой целочисленной переменной x_{ij} поставим в соответствие выходной сигнал ij -ого нейрона Y_{ij} , стоящего в i -ой строке и j -том столбце матрицы сети, т.е. $x_{ij} = 1 \Leftrightarrow Y_{ij} = 1$, при $\forall i \in 1 \dots N$ и $\forall j \in 1 \dots M$.

Совокупность возбужденных нейронов будем интерпретировать как схему подключения, тогда ограничение (2), можно рассмотреть как:

$$\sum_{j=1}^M Y_{ij} = 1, \quad \forall j = 1 \dots M, \quad (2')$$

т.е. хост может быть подключен только к одному коммутатору, а ограничение (3) как:

$$\sum_{i=1}^N \sum_{j=1}^M Y_{ij} = N, \quad (3')$$

где N - количество хостов, и ограничение (4) как:

$$\sum_{i=1}^N \sum_{j=1}^M B_{ij} Y_{ij} = \sum_{j=1}^M P_{jj+1} \quad (4')$$

Тогда целевую функцию можно представить в виде:

$$Q = \sum_{i=1}^N \sum_{j=1}^M Y_{ij} m_{ij} \quad (5')$$

Сконструируем энергетическую функцию $E(Y)$, минимизация которой обеспечивает выполнение ограничений (2-4) и решение задачи (6).

Такая функция будет содержать два компонента [9], один из которых обеспечивает выполнение ограничений, а второй минимизацию целевой функции:

$$E(Y) = E_L(Y) + E_Q(Y), \quad (7)$$

где:

$$E_Q(Y) = F \sum_{i=1}^N \sum_{j=1}^M Y_{ij} m_{ij}, \quad (8)$$

$$E_L(Y) = \frac{G}{2} \sum_{i=1}^N \sum_{j=1}^M \sum_{v \neq j} Y_{ij} Y_{iv} + \frac{B}{2} \left(\sum_{i=1}^N \sum_{j=1}^M Y_{ij} B_{ij} - \sum_{j=1}^M P_{jj+1} \right)^2 + \frac{C}{2} \left(\sum_{i=1}^N \sum_{j=1}^M Y_{ij} - n \right)^2 \quad (9)$$

G, H, C, F – положительные константы, подбираемые эмпирически.

Первое слагаемое принимает минимальное и равное нулю значение лишь в том случае, если каждая строка матрицы подключений содержит не более одной единицы, второе слагаемое принимает минимальное и равное нулю значение лишь в том случае, если суммарная интенсивность потока сообщений подключенных к коммутатору хостов не превышает допустимой пропускной способности канала связи между двумя соседними коммутаторами. Третье слагаемое принимает минимальное значение, если во всей матрице подключений содержится ровно N единиц (все N хостов подключены).

Просуммировав выражения (23) и (24) можно получить энергетическую функцию в завершеном виде:

$$E(Y) = F \sum_{i=1}^N \sum_{j=1}^M Y_{ij} m_{ij} + \frac{G}{2} \sum_{i=1}^N \sum_{j=1}^M \sum_{v \neq j} Y_{ij} Y_{iv} + \frac{B}{2} \left(\sum_{i=1}^N \sum_{j=1}^M Y_{ij} B_{ij} - \sum_{j=1}^M P_{jj+1} \right)^2 + \frac{C}{2} \left(\sum_{i=1}^N \sum_{j=1}^M Y_{ij} - n \right)^2 \quad (10)$$

Таким образом, построенная функция энергии достигает своего минимума во всех состояниях удовлетворяющих совокупности ограничений (2)-(4) и целевой функции (5).

Для того чтобы задать начальное состояние сети, сопоставим полученную функцию энергии с энергетической функцией, записанной в общем виде:

$$E(Y, W, I) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^M \sum_{\mu=1}^N \sum_{\nu=1}^M W_{ij\mu\nu} Y_{ij} Y_{\mu\nu} + \sum_{i=1}^N \sum_{j=1}^M Y_{ij} I_{ij} \quad (11)$$

где $W_{ij\mu\nu}$ - значение синапса между входом ij -ого нейрона и выходом $\mu\nu$,

I_{ij} - смещение ij -ого нейрона.

Для определения параметров сети в соответствии с построенной энергетической функцией (10), приведем ее к виду (11):

$$E(Y) = \frac{G}{2} \sum_{i=1}^N \sum_{j=1}^M \sum_{v \neq j} Y_{ij} Y_{iv} + \frac{H}{2} \sum_{i=1}^N \sum_{j=1}^M \sum_{\mu=1}^N \sum_{\nu=1}^M Y_{ij} Y_{\mu\nu} B_{ij} B_{\mu\nu} - H \sum_{j=1}^M P_{jj+1} \sum_{i=1}^N \sum_{j=1}^M Y_{ij} B_{ij} + \frac{C}{2} \sum_{i=1}^N \sum_{j=1}^M \sum_{\mu=1}^N \sum_{\nu=1}^M Y_{ij} Y_{\mu\nu} - C_n \sum_{i=1}^N \sum_{j=1}^M Y_{ij} + \frac{H}{2} \left(\sum_{j=1}^M P_{jj+1} \right)^2 + \frac{C}{2} n^2 \quad (12)$$

Сопоставление линейных членов позволит получить значения внешних смещений, а квадратичные члены помогут определить синаптические веса между нейронами [9]. Анализ первого слагаемого свидетельствует о том, что любой нейрон должен иметь связи с коэффициентом - G со всеми нейронами одноименной с ним строки (условие $\mu = i$) кроме самого рассматриваемого нейрона (условие $\nu \neq j$). Второе слагаемое обуславливает наличие связей с

коэффициентом $H \sum_{i=1}^N \sum_{j=1}^M \sum_{\mu=1}^N \sum_{\nu=1}^M B_{ij} B_{\mu\nu}$ между нейронами одного столбца (условие $\nu = j$) кроме собственной обратной связи (условие $\mu \neq i$).

Четвертое слагаемое определяет связи всех нейронов сети друг с другом синапсами с коэффициентами C .

Таким образом, веса синапсов будут определяться следующим образом:

$$W_{ij\mu\nu} = -G \delta_{i\mu} (1 - \delta_{i\mu}) - H \sum_{i=1}^N \sum_{j=1}^M \sum_{\mu=1}^N \sum_{\nu=1}^M B_{ij} B_{\mu\nu} \delta_{j\nu} (1 - \delta_{i\mu}) - c \quad (13)$$

где $\delta_{ij} = 1$, если $i = j$, и $\delta_{ij} = 0$, если $i \neq j$ - символ Кронекера.

Третье, пятое и шестое слагаемые определяют внешнее смещение:

$$I_{ij} = -\frac{H}{2} \left(\sum_{j=1}^M P_{jj+1} \right)^2 - C_n - F m_{ij} \quad (14)$$

В полученных выражениях значения пропускных способностей каналов связи B_{ij} и взвешенных длин кабельных соединений нормированы. Следовательно, получив начальные значения параметров нейросетевой модели можно представить ее обобщенную структуру.

Представленная модель позволяет представить задачу определения топологии СПД в виде пошагового алгоритма, при этом в матрице значений Y_{ij} , возбужденные нейроны (имеющие единичные выходы) каждого столбца, означают количество подключенных хостов к соответствующему коммутатору.

В завершении доклада приведём работы авторского коллектива в развитие указанного подхода [10-20].

Литература

1. Рожнов А.В., Лобанов И.А., Бимаков Е.В. Обоснование задач системной интеграции и информационно-аналитическое моделирование проблемно-ориентированных системы управления на предпроектном этапе жизненного цикла / Труды XII Всероссийского совещания по проблемам управления (ВСПУ-2014, Москва). М.: ИПУ РАН, 2014. С. 7474-7479.
2. Лобанов И.А., Рожнов А.В. Оценивание эффективности проблемно-ориентированной системы управления на ранних стадиях жизненного цикла комплекса ЛА с использованием модели FREE DISPOSAL HULL / Материалы 5-й Международной научной конференции "Фундаментальные проблемы системной безопасности и устойчивости" (Елец, 2014). Елец: Елецкий государственный университет им. И.А. Бунина, 2014. С. 377-379.
3. Антиох Г.М., Рожнов А.В. Исследование задач интегрированной системы моделирования и анализа эффективности среды функционирования в развитии идеи Joint Warfare System / Труды 14-й Международной конференции «Системы проектирования, технологической подготовки производства и управления этапами жизненного цикла промышленного продукта» (CAD/CAM/PDM-2014, Москва). М.: ООО "Аналитик", 2014. С. 101-103.
4. Поляк Б. Т. Введение в оптимизацию. – М.: Наука. – 1983. – 132 с.
5. Хетагуров Я. А. Основы проектирования управляющих вычислительных систем.– М.: Радио и связь.– 1991.– 288 с.
6. Понтрягин Л.С. Основы комбинаторной топологии. – М.: Наука. – 1986. – 263 с.
7. Амосов Н. М. и др. Нейрокомпьютеры и интеллектуальные роботы. – Киев: Наук. Думка. – 1991. – 272 с.
8. Головкин В. А. Нейронные сети: обучение организация и применение. Кн.4: Учеб. пособие для вузов / Общая ред. А.И. Галушкина. – М.:ИПРЖ. – 2001. – 256 с.
9. Ефимов В. В. Нейроподобные сети в бортовых информационно-управляющих комплексах космических аппаратов. – СПб.: ВИКА им. А.Ф. Можайского. – 1996. – 113 с.
10. Оганджян С.Б., Рожнов А.В., Бурмистров П.А., Лобанов И.А., Тюрин С.А. Творческие материалы «круглого стола». Часть I. Ретроспектива и реальная конкорданция исследований в сфере интеллекта // Нейрокомпьютеры: разработка, применение. 2016. № 1. С. 17-29.
11. Рожнов А.В. Творческие материалы «круглого стола». Часть II. Системная интеграция и моделирование новых эффектов в сфере интеллекта // Нейрокомпьютеры: разработка, применение. 2016. № 3. С. 3-12.
12. Рожнов А.В. Некоторые проблемные вопросы системной интеграции направлений научной деятельности в задачах нейрокомпьютинга // Нейрокомпьютеры: разработка, применение. 2014. № 1. С. 3-9.
13. Русаков К.Д., Гончаренко В.И., Горелов А.В. Модель прогноза технического состояния космического аппарата на основе искусственных модульных нейросетей / Труды 19-й Международной конференции «Цифровая обработка сигналов и ее применение - DSPA-2017» (Москва, 2017). М.: РНТОРЭС имени А.С. Попова, 2017. Т. 2. С. 808-813.
14. Рожнов А.В., Кривоножко В.Е., Лычев А.В. Построение гибридных интеллектуальных информационных сред и компонентов экспертных систем на основе обобщенной модели анализа среды функционирования // Нейрокомпьютеры: разработка, применение. 2013. № 6. С. 3-12.
15. Оганджян С.Б., Рожнов А.В., Гудова Л.П. Актуализация и особенности понимания в русскоязычной целевой аудитории общей дефиниции «искусственный интеллект» / Тезисы докладов 15-й Всероссийской научной конференции "Нейрокомпьютеры и их применение" (Москва, 2017). М.: ФГБОУ ВО МГППУ, 2017. С. 232-233.
16. Прокопенков А.А., Рожнов А.В., Карпов В.В. Управление сложными системами и конвергентные когнитивно-информационные технологии (введение/ Тезисы докладов 15-й Всероссийской научной конференцией "Нейрокомпьютеры и их применение" (Москва, 2017). М.: ФГБОУ ВО МГППУ, 2017. С. 268-270.
17. Легович Ю.С., Рожнов А.В., Руженцев И.О. Интеллектуальные права и их защита при разработке прорывных технологий и компонентов гибридных интеллектуальных и интеллектуализированных систем / Тезисы докладов 15-й Всероссийской научной конференцией "Нейрокомпьютеры и их применение" (Москва, 2017). М.: ФГБОУ ВО МГППУ, 2017. С. 49-50.
18. Лобанов И.А., Скорик Н.А. Формирование виртуальной семантической среды для управления смешанными робототехническими группировками / Тезисы докладов XIV Всероссийской научной конференции «Нейрокомпьютеры и их применение» (Москва, 2016). М.: ГБОУ ВПО МГППУ, 2016. С. 29-31.
19. Лобанов И.А. Регенеративный анализ в задаче формирования адаптивного информационно-управляющего пространства / Тезисы докладов XIV Всероссийской научной конференции «Нейрокомпьютеры и их применение» (Москва, 2016). М.: МГППУ, 2016. С. 26-28.
20. Волков А.Г., Полоус А.И., Гончаренко В.И., Иконников М.А. Синтез алгоритма вычисления оценок, адаптированного к нейросетевому логическому базису // Нейрокомпьютеры: разработка, применение. 2017. № 3. С. 34-41.