

Санкт-Петербургский государственный университет

Прикладная математика, программирование и искусственный интеллект

Отчет по учебной практике 2 (научно-исследовательской работе) (семестр 4)
Правила графов для динамики рекуррентных нейронных сетей

Выполнил:

Глумакова Дарья Андреевна, группа

23.Б06-мм



Научный руководитель:

Профессор кафедры прикладной
кибернетики, доктор физ.-мат. наук,

Мокаев Тимур Назирович

*Работа выполнена полностью
и может быть зачтена с оценкой А.*



Санкт-Петербург

2025

ОТЗЫВ научного руководителя

Тема работы «Правила графов для динамики рекуррентных нейронных сетей» является актуальной. При выполнении научно-исследовательской работы студентка Глумакова Дарья Андреевна проявила самостоятельность и целеустремленность в изучении новой области и получении результатов. Сильные стороны работы:

1. **Актуальность темы.** Исследуются современные методы анализа динамики рекуррентных нейронных сетей, имеющие непосредственное применение в вычислительной нейронауке и ИИ.
2. **Глубина теоретического анализа.** Автор систематизировала шесть элементарных правил графовой теории, вывела обобщения (правила равномерной степени, доминирования, склеивания) и корректно доказала ключевые леммы и теоремы.
3. **Высокая степень самостоятельности.** Проведена обширная работа с первоисточниками.
4. **Качество результатов.** Чётко показана связь структурных свойств графа с устойчивыми и неустойчивыми фиксированными точками CTLN-сетей; приведены наглядные иллюстрации и примеры, подтверждающие выводы.

Работа выполнена на хорошем уровне, отчет составлен грамотно и достаточно полно отражает результаты работы студента. Есть небольшие замечания к оформлению и срокам сдачи работы. В процессе обучения освоены компетенции **ОПК-3, ОПК-4, ПКА-2, ПКП-1, ПКП-2, ПКП-3, ПКП-4, УК-2, УКБ-1, УКБ-2**. Работа выполнена полностью и может быть зачтена с оценкой А.

Профессор кафедры прикладной кибернетики
д.ф.-м.н. Мокаев Т.Н.



9 июня 2025 г.

Оглавление

1. Введение.....	3
2. Цель и задачи.....	4
3. Рекуррентные нейронные сети	4
3.1. TLN.....	6
3.2. CTLN	7
4. Фиксированные точки.....	7
5. Правила графов	9
5.1. Основные правила	9
5.2. Правила равномерной степени	10
5.3. Графовое доминирование	12
5.4. Просто-вложенные подграфы.....	14
5.5. Элементарные правила графов.....	16
6. Правила склеивания	18
6.1. Связные объединения.....	19
6.2. Просто-вложенные разбиения	20
6.3. Основные мотивы (core motifs)	22
7. Заключение	25
8. Список литературы	26

1. Введение

Говоря о нейронных сетях стоит обратиться к естественным сетям в мозгу. Такие нейроны могут возбуждаться как самопроизвольно, так и в ответ на внешние стимулы. Кроме того, они обладают положительной обратной связью, которая может стать причиной возникновения неконтролируемого возбуждения в сети и привести к приступу эпилепсии, повышенной тревожности, галлюцинациям и иным нарушениям. Для предотвращения этого в реальных нейронных сетях преобладает ингибирование (подавление) — оно не только помогает остановить чрезмерное возбуждение нейронов, но и способствует формированию сложных паттернов активности, таких как ритмы, последовательности или переключение внимания. Это работает благодаря существованию различных нейронов и их сигналов. Нейроны взаимодействуют друг с другом с помощью электрических сигналов, называемых потенциалами действия. Эти сигналы могут быть возбуждающими, побуждая нейрон передавать информацию следующему, или тормозящими, снижающими вероятность передачи сигнала. Взаимодействие этих токов и их сбалансированность восстанавливают баланс системы и стабилизируют ее работу.

Один из главных вопросов в нейронауке заключается в том, возможно ли изучать и анализировать нейронную сеть основываясь только на ее структуре. Для ответа на этот вопрос разрабатываются теории, которые могут связать нелинейные динамические системы с их графовым представлением. С их помощью модели нейронных сетей, представленные в виде графов, позволяют предсказывать поведение сети и определять устойчивые и неустойчивые фиксированные точки, опираясь только на структуру графа. Фиксированные точки играют важную роль в моделировании поведения сети: устойчивые фиксированные точки представляют собой состояния памяти, а неустойчивые способствуют возникновению динамических аттракторов, таких как циклы.

Этот метод работает как для анализа простых сетей, так и для более сложных архитектур, состоящих из меньших подграфов, так как для них разработаны правила склеивания. Эти правила позволяют объединять фиксированные точки отдельных подграфов, чтобы получить полное представление о фиксированных точках всей сети.

Такой подход открывает новые возможности для анализа динамики нейросетей, позволяя исследовать сложные системы с использованием информации только о структуре графа, без необходимости анализировать каждую связь в отдельности. [2,3]

2. Цель и задачи исследования

Цель данной работы – изучение динамики рекуррентных нейронных сетей, в частности рассматривая порогово-линейные и комбинаторные порогово-линейные искусственные нейронные сети. В рамках исследования были поставлены следующие задачи:

1. Исследовать рекуррентных нейронных сетей, включая порогово-линейные (TLNs) и комбинаторные порогово-линейные нейронные сети (CTLNs).
2. Проанализировать понятия фиксированных точек, их классификации и роли в динамике нейронных сетей.
3. Исследовать графовые правила, которые позволяют анализировать динамику нейронных сетей, основываясь исключительно на графовом представлении их структуры.
4. Изучить правила склейки, которые позволяют комбинировать более простые сети в более сложные и анализировать их стабильные и нестабильные фиксированные точки.

Таким образом, основная задача работы заключается в теоретическом исследовании фиксированных точек, их роли в динамике нейронных сетей и применении графовых и склеивающих правил для анализа сетей.

3. Рекуррентные нейронные сети

Рекуррентная нейронная сеть (RNN) – это искусственная нейронная сеть, в которой связи между нейронами образуют обратные связи: выход нейрона на одном шаге времени возвращается в сеть и дополнительно подается на следующий вход. Благодаря этому RNN может запоминать информацию о предыдущих шагах. В поставленной задаче рекуррентные сети в первую очередь рассматриваются как динамические системы, основная цель которых - моделировать динамику нейронной активности в головном мозге. Простая схема работы RNN представлена на Рис.1.

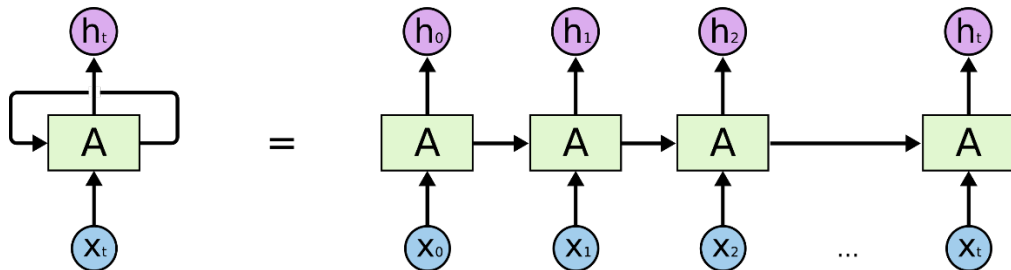


Рис.1. Развернутое представление RNN [1]

Так, RNN может быть представлена в виде ориентированного графа, в котором вершины представляют нейроны, а ребра – связи между нейронами. Каждой вершине сопоставляется функция $x_i(t)$, которая отслеживает уровень активности нейрона i во времени.

Кроме того, каждой паре нейронов (i, j) присваивается вес W_{ij} , который определяет силу влияния нейрона j на нейрон i . Это влияние осуществляется через обратные связи: выход нейрона j на предыдущем шаге времени $x_j(t - 1)$ поступает в нейрон i на текущем шаге и, с учётом соответствующего веса W_{ij} , влияет на его активность $x_i(t)$. Эти веса являются ключевыми параметрами сети, которые могут изменяться в процессе обучения для оптимизации работы модели. [2]

Динамика активности нейронов часто описывается системой обыкновенных дифференциальных уравнений:

$$\tau_i \frac{dx_i}{dt} = -x_i + \varphi \left(\sum_{j=1}^n W_{ij} x_j + b_i \right) = -x_i + \varphi(y_i)$$

Где:

- x_i — уровень активности нейрона i ,
- y_i — общий входной сигнал нейрона i ,
- τ_i — характеристическое время затухания активности нейрона i ,
- b_i — внешний вход в нейрон i ,
- φ — функция активации, обычно монотонно возрастающая.

3.1. TLN

Особое значение для предмета изучения представляют порогово-линейные нейронные сети (TLN). В таких сетях в качестве функции активации используется порогово-линейная функция (ReLU), представимая следующей формулой:

$$\varphi(y) = [y]_+ = \max \{0, y\}$$

То есть нейрон активируется только когда сумма входных сигналов превышает некоторое пороговое значение (здесь 0). Такая форма нелинейности делает такие модели удобными для анализа.

Мы так же полагаем, что $W_{ii} = 0$ для каждого i . Тогда при постоянной матрице весов W и входном векторе $b = (b_1, \dots, b_n)$ вход каждого нейрона определяется суммой сигналов от других нейронов с добавлением внешнего входа:

$$y_i = \sum_{j=1}^n W_{ij}x_j + b_i = 0$$

Каждое такое выражение можно интерпретировать как линейное уравнение, задающее гиперплоскость в пространстве \mathbb{R}^n . Каждая такая гиперплоскость H_i определяется уравнением

$$y_i = n_i * x + b_i = 0$$

Где:

- $n_i = (W_{i1}, \dots, W_{in})$ – вектор нормали
- $x = (x_1, \dots, x_n)$ – вектор активности нейронов
- b_i – смещение

Совокупность всех таких гиперплоскостей $H = \{H_1, \dots, H_n\}$ делят положительный ортант $\mathbb{R}_{\geq 0}^n$ на области, которые называются камерами. Внутри каждой камеры каждое состояние x находится с положительной или отрицательной стороны каждой гиперплоскости H_i в зависимости от знака y_i . Уравнения системы тогда сводятся к линейной системе дифференциальных уравнений, где для каждого i выполняется одно из двух уравнений:

$$\begin{aligned} \frac{dx_i}{dt} &= -x_i + y_i, \text{ если } y_i > 0 \\ \frac{dx_i}{dt} &= -x_i, \text{ если } y_i \leq 0 \end{aligned}$$

Таким образом гиперплоскости помогают анализировать переходы между различными состояниями системы и её поведение в различных областях пространства состояний.

Вспомним, что в реальных нейронных сетях ингибирование играет ключевую роль в стабилизации активности сети, так как без ингибиторных связей нейроны могли бы разогнаться в неконтролируемой активности. Ингибирование помогает поддерживать баланс в сети. Учитывая это становится понятно, что для нашей задачи важно изучать ингибиторные TLN, то есть такие, в невырожденной матрице весов W которых все элементы неположительны ($W_{ij} \leq 0 \forall ij$).

Такие сети естественным образом моделируют процессы выбора между несколькими альтернативами, где "побеждает" тот нейрон который сильнее подавляет других. Эти свойства делают ингибиторные сети особенно важными в нейронауке для моделирования процессов восприятия и памяти. [2]

3.2. CTLN

Для дальнейшей работы наибольший интерес представляет особый класс ингибиторных TLNs – комбинаторные порогово-линейные нейронные сети (CTLNs). Эти модели фокусируются исключительно на возбуждающих связях между нейронами, в то время как ингибирование реализуется глобально и равномерно.

Ингибиторный эффект реализуется через глобальный входной сигнал – постоянный тормозной вклад, одинаковый для всех нейронов. Это создает баланс возбуждения и ингибирования в сети.

Динамика такой сети практически полностью определяется структурой графа G , в котором каждая вершина соответствует нейрону, а каждое направленное ребро – связи между нейронами. При этом веса связей W_{ij} принимают значения в зависимости от наличия или отсутствия ребра в графе:

$$W_{ji} = \begin{cases} 0, & \text{если } i = j \\ -1 + \varepsilon, & \text{если } j \rightarrow i \\ -1 - \delta, & \text{если } j \nrightarrow i \end{cases}$$

Где $\varepsilon > 0$, $\delta > 0$, и $\varepsilon < \frac{\delta}{\delta+1}$. Это ограничение обеспечивает, что торможение всегда будет доминировать и модель остается ингибиторной, так как все $W_{ij} \leq 0$.

Кроме того, в CTLN обычно используется постоянный входной сигнал $b_i = \theta$, $\theta > 0$, что гарантирует, что динамика сети зависит только от ее внутренних связей. Так, поведение комбинаторно-пороговой нейронной сети целиком задается тремя параметрами $(\varepsilon, \delta, \theta)$. Это делает модель удобной для анализа, так как позволяет изучать динамику только на основе структуры графа без необходимости подбирать или обучать веса.[2]

4. Фиксированные точки

В динамических системах фиксированная точка — это состояние в котором система перестает изменяться со временем. Для порогово-линейных нейронных сетей и комбинаторных порогово-линейных нейронных сетей фиксированные точки играют ключевую роль в анализе их поведения, так как они определяют устойчивость системы, переходы между состояниями, возникновение сложных динамик и аттракторов.

Фиксированная точка для системы – это вектор $x^* \in \mathbb{R}^n$, в котором все производные от состояний нейронов равны нулю, т.е.

$$\frac{dx_i}{dt} \big|_{x=x^*} = 0 \text{ для } \forall i \in \{1, \dots, n\}$$

В случае TLN это эквивалентно $x_i^* = [y_i^*]_+$ для $\forall i \in \{1, \dots, n\}$, где $[y_i^*]_+$ – пороговая функция активации.

Таким образом, в фиксированной точке активность каждого нейрона не изменяется во времени.

Существует два типа фиксированных точек:

1. Стабильная фиксированная точка — это точка, в которой система стабилизируется после небольших отклонений. Они являются статическими аттракторами.
2. Нестабильная фиксированная точка — это точка, в которой любые отклонения от состояния равновесия приводят к дальнейшему уходу от неё. В таких случаях

система может перейти в другое состояние, например, начать колебания или перейти в режим доминирования другого нейрона. Эти точки играют важную роль в формировании динамических аттракторов, например, таких как предельные циклы.

На Рис.2. представлены стабильные и нестабильные фиксированные точки. Стабильная фиксированная точка (А) служит аттрактором сети — при любом малом отклонении система стремится вернуться к ней. Нестабильные фиксированные точки (В, С), напротив, не являются аттракторами, однако некоторые из них (В) связаны с динамическими аттракторами, определяя последовательность активации нейронов, а другие (С) работают только как переломные точки между аттракторами.



Рис.2. Фиксированные точки. Figma. [2]

Для того, чтобы классифицировать фиксированные точки сети необходимо рассмотреть собственные значения матрицы $J_\sigma = -I + W_\sigma$:

1. Если для $\forall i \operatorname{Re}(\lambda_i) < 0$, то фиксированная точка устойчива.
2. Если $\exists i \operatorname{Re}(\lambda_i) > 0$, то фиксированная точка является неустойчивой.
3. Если $\operatorname{Re}(\lambda_i) = 0$, то случай требует отдельного анализа. [2]

Состояние фиксированной точки $x^* \in \mathbb{R}^n$ определяется ее поддержкой. Она представляет собой подмножество активных нейронов, то есть таких, которые имеют положительные значения в фиксированной точке:

$$\operatorname{supp} x^* = \{i \mid x_i^* > 0\}$$

Невырожденность матрицы весов гарантирует, что в каждой камере гиперплоскости $H(W, b)$ может быть не более одной фиксированной точки. Таким образом, для каждого множества поддержек можно выделить по одной фиксированной точке. Множество всех фиксированных точек для заданной сети можно обозначить следующим образом:

$$FP(W, b) = \{\sigma \subseteq [n] \mid \sigma = \operatorname{supp} x^*\}$$

для некоторой фиксированной точки x^* , где $[n] = \{1, \dots, n\}$.

Это множество можно частично упорядочить по включению. Минимальные элементы — это такие поддержки, которые не содержат в себе другие поддержки фиксированных точек.

Каждое такое поддерживающее множество $\sigma \in FP(W, b)$ позволяет восстановить фиксированную точку. Вне поддержки, для всех индексов $i \notin \sigma$, значения $x_i^* = 0$, а внутри поддержки фиксированная точка x_σ^* определяется следующим образом:

$$x_\sigma^* = (I - W_\sigma)^{-1} * b_\sigma$$

Где x_σ^* и b_σ это векторы, полученные сужением векторов x и b системы на нейроны, входящие в поддержку σ , а матрица W_σ – подматрица полученная путем ограничения строк и столбцов матрицы весов W на нейроны из σ .

Важное свойство фиксированной точки x^* с поддержкой σ в том, что должны выполняться следующие условия:

- Для $\forall i \in \sigma : y_i^* > 0 \Rightarrow x_i^* = y_i^*$
- Для $\forall k \notin \sigma : y_k^* \leq 0 \Rightarrow x_k^* = 0$

Следует отметить, что если поддержка σ является фиксированной точкой в подсети (W_σ, b_σ) , то она также будет фиксированной точкой в полной сети (W, b) , но обратное не всегда верно, так как условия для нейронов, не входящих в поддержку, могут нарушаться в большей сети.

Теперь введем понятие основного мотива. Индуцированный σ подграф $G - G|_\sigma$ называется основным мотивом, если он имеет ровно одну фиксированную точку, и эта фиксированная точка имеет полную поддержку, то есть в фиксированной точке все нейроны в σ активны:

$$FP(G|_\sigma) = \{\sigma\}$$

Если такой мотив σ сохраняется как фиксированная точка в более крупной CTLN - сети G , то он называется сохранившимся основным мотивом, и он предсказывает наличие аттрактора.

5. Правила графов

5.1. Основные правила

Чтобы установить, какие подмножества $\sigma \subseteq [n]$ могут быть поддержками фиксированных точек, удобно использовать представление нейронной сети в виде ориентированного графа. Перед формулировкой правил введём ключевые понятия из теории графов:

1. Источник(source) – вершина без входящих ребер
2. Сток(sink) – вершина без исходящих ребер
3. Независимое множество – набор узлов без ребер между ними
4. Клика – набор вершин, любые две из которых соединены ребром.

5. Цикл – это граф или подграф, в котором каждый узел имеет ровно одно входящее и одно исходящее ребра.
6. Ориентированный ациклический граф(DAG) – это граф или подграф с такими топологически упорядоченными вершинами, что $i \rightarrow j$ при $\forall i > j$. Такой граф не содержит ориентированных циклов.
7. Цель графа $G|_{\sigma}$ - узел k , такой, что для всех $i \in \sigma \setminus \{k\}$ $i \rightarrow k$. Цель может находиться как внутри, так и вне графа.

В таблице ниже представлены восемь основных правил, иллюстрирующих, как свойства подграфов влияют на принадлежность подмножеств σ к $FP(G)$. [2]

№	Структура $G _{\sigma}$	Правило
1	Независимое множество	Если $G _{\sigma}$ – независимое множество то $\sigma \in FP(G)$ тогда и только тогда, когда каждая вершина $i \in \sigma$ является стоком в графе G (т.е. не имеет исходящих рёбер).
2	Клика	Если $G _{\sigma}$ – клика, то $\sigma \in FP(G)$ тогда и только тогда, когда нет вершины $k \in G, k \notin \sigma$, т.ч. $i \rightarrow k \forall i \in \sigma$. Если $\sigma \in FP(G)$, то фиксированная точка будет устойчивой.
3	Циклы	Если $G _{\sigma}$ – цикл, то $\sigma \in FP(G)$ тогда и только тогда, когда нет вершины $k \in G, k \notin \sigma$, в которую входят два или более ребер из σ . Если $\sigma \in FP(G)$, то фиксированная точка будет неустойчивой.
4(i)	Источник	Если в $G _{\sigma}$ есть вершина $j \in \sigma$, являющаяся источником, и существует $j \rightarrow k$, для некоторого $k \in [n]$ то $\sigma \notin FP(G)$
4(ii)	Источник	Если $j \notin \sigma$, но j – источник в G , тогда $\sigma \in FP(G _{\sigma \cup j}) \Leftrightarrow \sigma \in FP(G _{\sigma})$.
5(i)	Цель	Если в σ есть вершина k , являющаяся целью (все остальные вершины из σ направлены в неё), и при этом $k \rightarrow j$ для некоторого $j \in \sigma$ ($j \neq k$), то $\sigma \notin FP(G _{\sigma \cup k})$, и соответственно, $\sigma \notin FP(G)$.
5(ii)	Цель	Если $k \rightarrow \sigma$, и k — цель σ , то $\sigma \notin FP(G _{\sigma \cup k})$, и значит, $\sigma \notin FP(G)$.
6	Сток вне поддержки	Если в G есть сток $s \notin \sigma$, то $\sigma \cup \{s\} \in FP(G) \Leftrightarrow \sigma \in FP(G)$
7	Ациклический граф	Если G – направленный ациклический граф со стоками s_1, \dots, s_{ℓ} , то $FP(G) = \{\cup s_i s_i \text{ – сток в } G\} = 2^{\ell} - 1$ элементов.
8	Четность	Для любого графа G количество поддержек фиксированных точек $ FP(G) $ всегда нечётно.

Таб.1. Правила графов. [2]

5.2. Правила равномерной степени

Помимо представленных восьми правил, существует более общий набор условий, объясняющих, какие подмножества нейронов могут поддерживать фиксированные точки. Так, правила 1, 2 и 3 являются следствиями правила для графов равномерной степени.

Граф $G|_{\sigma}$ называется графом с равномерной входящей степенью d если каждая вершина $i \in \sigma$ имеет ровно d входящих ребер от других вершин из σ . Примеры графов с указанной равномерной степенью представлены на Рис. 3.

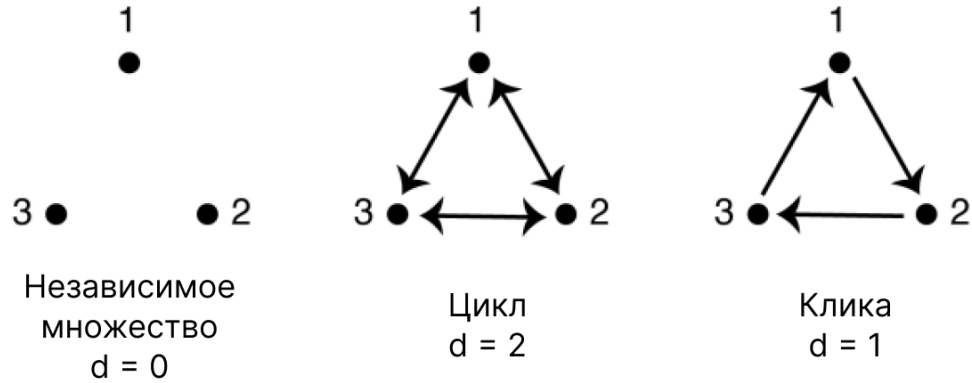


Рис.3. Входящая степень.Figma. [2]

Для CTLN фиксированная точка x^* поддерживаемая множеством σ удовлетворяет уравнению:

$$(I - W_{\sigma})x_{\sigma}^* = \theta * 1_{\sigma}$$

Где:

- W_{σ} – подматрица весов между узлами σ
- θ – порог
- 1_{σ} – вектор из единиц размерности $|\sigma|$ (собственный вектор $I - W_{\sigma}$)

Если граф $G|_{\sigma}$ имеет равномерную входящую степень d , то все строки матрицы $I - W_{\sigma}$ имеют одинаковую сумму. Каждая ее строка соответствует одному активному нейрону $i \in \sigma$. Обозначим их сумму через R . Тогда решение принимает вид:

$$x_{\sigma}^* = \frac{\theta}{R} 1_{\sigma}$$

Для входной степени d сумма R может быть вычислена следующим образом:

$$R = 1 + d(1 - \varepsilon) + (|\sigma| - d - 1)(1 + \delta)$$

Где:

- 1 – элемент с диагонали
- $d * (-(-1 + \varepsilon)) = d * (1 - \varepsilon)$ – вклад d входящих нейронов
- $(|\sigma| - d - 1) * (-(-1 - \delta)) = (|\sigma| - d - 1) * (1 + \delta)$ – вклад остальных нейронов

Зная это, можно привести формулу для значения фиксированной точки к следующему виду:

$$x_i^* = \frac{\theta}{|\sigma| + \delta(|\sigma| - d - 1) - \varepsilon d} 1_{\sigma}$$

Эта формула справедлива для любого графа с равной входящей степенью на σ , даже если граф не является симметричным. Таким образом, одинаковая входящая степень в графе

гарантирует, что в фиксированной точке все нейроны в нём будут иметь одинаковое значение их активации.[2]

Кроме того, с помощью этой формулы можно проверять выполнение условия активности нейронов, то есть что $x_i^* > 0$ для каждого $i \in \sigma$ при заданных параметрах $\theta, \delta, \varepsilon$. Для нейрона $k \notin \sigma$ значение правой части уравнения CTLN примет вид:

$$\begin{aligned} y_k^* &= \sum_{i \in k} W_{ki} x_i^* + \theta = \sum_{i \rightarrow k} (-1 + \varepsilon) x_i^* + \sum_{i \nrightarrow k} (-1 - \delta) x_i^* + \theta = \\ &= \theta * \left(\frac{d_{k(-1+\varepsilon)} + (|\sigma| - d_k)(-1 - \delta)}{|\sigma| + \delta(|\sigma| - d - 1) - \varepsilon d} \right) + 1 \end{aligned}$$

Где $d_k = |\{i \in \sigma | i \rightarrow k\}|$ – количество ребер из σ в k . Условие $y_k^* \leq 0$ выполнено тогда и только тогда, когда $d_k < d$.

Это приводит нас к следующей теореме:

Теорема 1

Пусть $G|_\sigma$ – индуцированный подграф G с равномерной входящей степенью d . Для каждого $k \notin \sigma$ обозначим через d_k количество ребер $i \rightarrow k$, где $i \in \sigma$. Тогда $\sigma \in FP(G|_\sigma)$ и $\sigma \in FP(G|_{\sigma \cup k}) \Leftrightarrow d_k \leq d$.

В частности, $\sigma \in FP(G) \Leftrightarrow \nexists k \notin \sigma$ такого, что $d_k > d$. [2]

Можно вывести следствие этой теоремы(частный случай):

Для этого рассмотрим множество $\sigma = [n]$. Тогда подграф $G|_\sigma = G$, и он, по условию, имеет равномерную входящую степень. Поскольку $\sigma = [n]$, то множество вершин $k \notin \sigma$ пусто. Условие теоремы выполняется тривиально: для всех $k \notin \sigma$ выполнено $d_k \leq d$, так как таких k просто нет. Следовательно $\sigma = [n] \in FP(G)$.

Следствие 1

Пусть G – граф с равномерной степенью с n вершинами. Тогда G имеет неподвижную точку $\sigma = [n] \in FP(G)$. [4]

5.3. Графовое доминирование

Как было показано ранее, в графах с равномерной входящей степенью фиксированные точки характеризуются одинаковым уровнем активности нейронов и было установлено важное ограничение: если существует $k \notin \sigma$, получающий слишком много входов от σ то σ не может быть фиксированной точкой. Однако существуют случаи когда подобный дисбаланс возникает внутри самого σ . Чтобы анализировать такие ситуации и исключать неверные неподвижные точки, введем понятие графового доминирования. [2, 4]

Пусть G – граф с n узлами, $\sigma \subseteq [n]$. Для $j, k \in [n]$ говорим, что k графово доминирует над j относительно σ , если выполняются следующие три условия:

1. Для любого $i \in \sigma \setminus \{j, k\}$, если $i \rightarrow j$ то $i \rightarrow k$.
2. Если $i \in \sigma$, то $j \rightarrow k$.
3. Если $k \notin \sigma$, то $k \nrightarrow j$.

В зависимости от того, находятся ли j и k внутри σ можно выделить 4 типа доминирования, которые показаны в Таблице 2.

Тип	Условия
Inside-in (оба внутри)	$j \in \sigma, k \in \sigma$
Outside-in (один снаружи, другой внутри)	$j \in \sigma, k \notin \sigma$
Inside-out (один внутри, другой снаружи)	$j \notin \sigma, k \in \sigma$
Outside-out (оба снаружи)	$j, k \notin \sigma$

Таб.2.

Для каждого из них можно вывести строгие утверждения о невозможности существования фиксированной точки. Рассмотрим их в следующей теореме:

Теорема 2

Если k графово доминирует над j относительно σ в G , то:

1. (inside-in) Если $j, k \in \sigma$, то $\sigma \notin G|_{\sigma}$ и следовательно $\sigma \notin FP(G)$.
Оба не могут быть активны одновременно: возникает противоречие в уравнениях на фиксированную точку ($y_j^* \leq y_k^*$ и $x_j^* > 0, x_k^* > 0$).
2. (outside-in) Если $j \in \sigma, k \notin \sigma$, то $\sigma \notin G|_{\sigma \cup k}$ и $\sigma \notin FP(G)$.
Поскольку j активен, k также должен быть активен, но это противоречит условию его неактивности.
3. (inside-out) Если $k \in \sigma, j \notin \sigma$, и $\sigma \in FP(G|_{\sigma})$, то $\sigma \in FP(G|_{\sigma \cup j})$.
Нейрон j гарантированно останется неактивным — его активность невозможна.
4. (outside-out) Если $j, k \notin \sigma$, и $\sigma \in FP(G|_{\sigma \cup k})$, то $\sigma \in FP(G|_{\sigma \cup j})$.
Если k неактивен, то и j не будет активным, так как у него ещё меньше входов. [2]

В основе доказательства лежит сравнение входов y_j^* и y_k^* , где:

$$y_j^* = \sum_{i \in \sigma \setminus \{j, k\}} W_{ji} x_i^* + W_{jk} x_k^* + \theta$$

$$y_k^* = \sum_{i \in \sigma \setminus \{j, k\}} W_{ki} x_i^* + W_{kj} x_j^* + \theta$$

а W_{ij} – матрица весов заданных следующим образом:

$$\begin{cases} W_{ij} = -1 + \varepsilon, \text{ если } i \rightarrow j \\ W_{ij} = -1 - \delta, \text{ если } i \nrightarrow j \end{cases}$$

При выполнении условий графового доминирования справедливо неравенство:

$$y_j^* + W_{kj} x_j^* \leq y_k^* + W_{jk} x_k^*$$

Это означает, что даже с учётом обратных влияний (внутри σ) нейрон k получает не меньший (а возможно и больший) сигнал, чем j . Это приводит к логическому противоречию в случаях inside-in и outside-in, а в inside-out и outside-out позволяет сделать выводы о невозможности активации j . [2, 4]

5.4. Просто-вложенные подграфы

Теперь введем еще одно важное понятие – просто-вложенный подграф:

Подграф $G|_\tau$ называется просто вложенным в граф G , если для $\forall k \notin \tau$ выполняется одно из двух условий:

- (i) $k \rightarrow i$ для всех $i \in \tau$
- (ii) $k \nrightarrow i$ для всех $i \in \tau$

То есть или все нейроны, не входящие в подграф направлены в узлы из τ , либо все в них не направлены. Тогда можем говорить, что подмножество вершин $\tau \subseteq [n]$ просто-вложено в G .

Рассмотрим как себя ведут нейроны из просто-вложенного подграфа. Для каждого нейрона $i \in \tau$, динамика описывается уравнением:

$$\frac{dx_i}{dt} = -x_i + \left[\sum_{j \in \tau} W_{ij}x_j + \sum_{k \notin \tau} W_{ik}x_k + \theta \right]_+$$

Где:

- $\sum_{j \in \tau} W_{ij}x_j$ – взвешенная сумма входов от нейронов изнутри τ
- $\sum_{k \notin \tau} W_{ik}x_k$ – взвешенная сумма входов от нейронов извне τ

Если подграф $G|_\tau$ просто-вложен, то все нейроны $i \in \tau$ получают одинаковое воздействие от нейронов вне τ . Это значит, что коэффициенты W_{ik} одинаковы для $\forall i \in \tau$, если $k \notin \tau$.

Значит можно ввести общее внешнее воздействие:

$$\mu_\tau(t) = \sum_{k \notin \tau} W_{ik}x_k(t) + \theta \text{ (для любого } i \in \tau \text{)}$$

Таким образом, уравнения для всех нейронов в τ можно записать следующим образом:

$$\frac{dx_i}{dt} = -x_i + \left[\sum_{j \in \tau} W_{ij}x_j + \mu_\tau(t) \right]_+$$

В фиксированной точке x_k^* значения всех x_k постоянны, а значит

$$\mu_\tau^* = \sum_{k \notin \tau} W_{ik}x_k^* + \theta = const$$

Когда подграф τ просто-вложен в сеть CTLN, все его нейроны получают одинаковый внешний вклад μ_τ^* от нейронов вне τ . Это позволяет рассматривать τ как отдельную CTLN-сеть с новым входным параметром $\theta' = \mu_\tau^*$.

1. Если $\theta' \leq 0$ то фиксированная точка в τ – нулевая, то есть поддержка $\sigma \cap \tau = \emptyset$

2. Если $\theta' > 0$, то $\sigma \cap \tau \in FP(G|_\tau)$

Это приводит к следующей лемме :

Лемма 1

Пусть $G|_\tau$ просто-вложен в G . Тогда для любого $\sigma \subseteq [n]$ верно:

$$\sigma \in FP(G) \Rightarrow \sigma \cap \tau \in FP(G|_\tau) \cup \{\emptyset\}$$

Рассмотрим теперь операции над просто вложенными подграфами. Пусть $\tau_1, \tau_2 \subseteq [n]$ – два просто вложенных подграфа. Операции над ними сохраняют просто-вложенность при определенных условиях, что рассмотрено в следующей лемме.

Лемма 2 (о пересечениях и объединениях)

Пусть $\tau_1, \tau_2 \subseteq [n]$ – просто-вложенные подмножества в G . Тогда:

1. $\tau_1 \cap \tau_2$ всегда просто-вложены
2. $\tau_1 \cup \tau_2$ просто-вложены, если $\tau_1 \cap \tau_2 = \emptyset$. [2]

Можем рассмотреть ее доказательство:

Если $\tau_1 \cap \tau_2 = \emptyset$, то пересечение является просто-вложенным. Предположим теперь, что $\tau_1 \cap \tau_2 \neq \emptyset$ и рассмотрим произвольную вершину $k \notin \tau_1 \cap \tau_2$. Если $k \notin \tau_1$, а τ_1 – просто-вложено то по определению просто-вложенности вершина k одинаково взаимодействует со всеми вершинами τ_1 , следовательно, и с подмножеством $\tau_1 \cap \tau_2$. Аналогично, если $k \notin \tau_2$, то она одинаково взаимодействует со всеми вершинами τ_2 и, следовательно, с $\tau_1 \cap \tau_2$. Таким образом, $\tau_1 \cap \tau_2$ – просто-вложено.

Рассмотрим теперь объединение $\tau_1 \cup \tau_2$ такие, что $\tau_1 \cap \tau_2 \neq \emptyset$. Пусть $j \in \tau_1 \cap \tau_2$ и $k \notin \tau_1 \cup \tau_2$. Если $k \rightarrow j$, то из определения просто-вложенности следует, что $k \rightarrow i$ для всех $i \in \tau_1$ и $k \rightarrow \ell$ для всех $\ell \in \tau_2$. В противном случае, если $k \nrightarrow j$, то аналогично $k \nrightarrow i$ для всех $i \in \tau_1$ и $k \nrightarrow \ell$ для всех $\ell \in \tau_2$. Следовательно, k одинаково взаимодействует со всеми вершинами $\tau_1 \cup \tau_2$, что доказывает просто-вложенность объединения. ■

Таким образом, операции пересечения и, при некоторых условиях, объединения сохраняют просто-вложенность подграфов. Таким образом это свойство позволяет строить более сложные структуры на их основе. Рассмотрим в частности как можно “склеить” просто вложенные графы и как при этом ведут себя их фиксированные точки. Для этого обратимся к следующей лемме:

Лемма 3

Считаем, что $\widehat{FP}(G) = FP(G) \cup \{\emptyset\}$.

Пусть $G|_{\tau_i}$ и $G|_{\tau_j}$ – просто-вложенные подграфы G . Рассматриваем $\sigma_i \in \widehat{FP}(G|_{\tau_i})$, $\sigma_j \in \widehat{FP}(G|_{\tau_j})$ такие, что $\sigma_i \cap \tau_j = \sigma_j \cap \tau_i$. Тогда $\sigma_i \cup \sigma_j \in \widehat{FP}(G|_{\tau_i \cup \tau_j})$ тогда и только тогда, когда выполняется одно из следующих условий:

- (i) $\tau_i \cap \tau_j = \emptyset$ и $\sigma_i, \sigma_j \in \widehat{FP}(G|_{\tau_i \cup \tau_j})$
- (ii) $\tau_i \cap \tau_j = \emptyset$ и $\sigma_i, \sigma_j \notin \widehat{FP}(G|_{\tau_i \cup \tau_j})$
- (iii) $\tau_i \cap \tau_j \neq \emptyset$

5.5. Элементарные правила графов

Наконец рассмотрим шесть элементарных правил, из которых могут быть выведены все остальные. Первые три правила справедливы для любых TLNs, остальные относятся к CTLNs и обобщают выводы из предыдущих пунктов[2]:

Правило 1

Это правило об уникальности поддержки фиксированной точки.

Для любого графа G существует не более одной фиксированной точки с заданной поддержкой $\sigma \subseteq [n]$.

Иначе говоря, если $\sigma \in FP(G)$, то существует единственная фиксированная точка x^* такая, что:

$$supp(x^*) = \sigma$$

Таким образом, фиксированные точки можно однозначно обозначить элементами множества $FP(G)$.

Правило 2

Пусть $\sigma \subseteq [n]$. Тогда $\sigma \in FP(G) \Leftrightarrow \sigma \in FP(G|_{\sigma})$ и $\sigma \in FP(G|_{\sigma \cup k})$ для всех $k \notin \sigma$.

То есть фиксированная точка с поддержкой σ существует в сети G тогда и только тогда, когда:

1. σ – фиксированная точка в индуцированном подграфе $G|_{\sigma}$,
2. Эта точка сохраняется при добавлении любого нейрона $k \notin \sigma$, то есть останется фиксированной точкой в расширенном графе.

Причём сохранится точка σ при добавлении вершины k или нет зависит только от рёбер $i \rightarrow k$, где $i \in \sigma$, и не зависит от рёбер в обратную сторону $k \rightarrow i$.

Правило 3

Общее количество фиксированных точек, $|FP(G)|$, всегда нечётно.

Это правило идентично правилу 8 из пункта 5.1.

Правило 4

В данном правиле опять обращаемся к равномерной степени:

$G|_{\sigma}$ – подграф с равномерной степенью. Тогда:

1. $\sigma \in FP(G|_{\sigma})$ – подграф сам по себе обладает фиксированной точкой.
2. $\sigma \in FP(G|_{\sigma \cup k}) \Leftrightarrow d_k \leq d$ в $G|_{\sigma \cup k}$ где d_k – количество рёбер, входящих в вершину k из σ .

Следовательно, $\sigma \in FP(G)$ тогда и только тогда, когда не существует вершины $k \notin \sigma$, получающей больше чем d входов от σ .

Это правило удобно для быстрой проверки, может ли подмножество нейронов σ быть поддержкой фиксированной точки.

Правило 5

Пусть вершина k доминирует над вершиной j относительно подмножества σ (в смысле графа: входящие и исходящие стрелки k «сильнее», чем у j). Тогда:

1. Доминирование inside-in: Если $j \in \sigma$ и $k \in \sigma$, то $\sigma \notin FP(G|_{\sigma})$, а значит $\sigma \notin FP(G)$.

2. Доминирование outside-in: Если $j \in \sigma$ и $k \notin \sigma$, то $\sigma \notin FP(G|_{\sigma \cup k})$, а значит $\sigma \notin FP(G)$.
3. Доминирование inside-out: Если $k \in \sigma, j \notin \sigma$, то из $\sigma \in FP(G|\sigma)$ следует $\sigma \in FP(G|_{\sigma \cup k})$.
4. Доминирование outside-out: Если $j \notin \sigma$ и $k \notin \sigma$, то из $\sigma \in FP(G|_{\sigma \cup k})$ следует $\sigma \in FP(G|_{\sigma \cup j})$.

Это правило позволяет исключать или включать фиксированные точки, сравнивая вершины по структуре ребер.

Правило 6

Пусть $G|_{\tau_i}$ и $G|_{\tau_j}$ — просто-вложенные подграфы графа G . Обозначим

$\widehat{FP}(G) = FP(G)$. Тогда:

1. Если $\sigma \in FP(G)$, то $\sigma \cap \tau_i \in \widehat{FP}(G|_{\tau_j})$.
2. Если $\sigma_i \in \widehat{FP}(G|_{\tau_j}), \sigma_j \in \widehat{FP}(G|_{\tau_i})$, и $\sigma_i \cap \tau_j = \sigma_j \cap \tau_i$. Тогда $\sigma = \sigma_i \cup \sigma_j \in \widehat{FP}(G|_{\tau_j \cup \tau_i})$ тогда и только тогда, когда выполняется одно из условий:
 - i. $\tau_i \cap \tau_j = \emptyset$ и $\sigma_i, \sigma_j \in \widehat{FP}(G|_{\tau_j \cup \tau_i})$,
 - ii. $\tau_i \cap \tau_j = \emptyset$ и $\sigma_i, \sigma_j \notin \widehat{FP}(G|_{\tau_j \cup \tau_i})$,
 - iii. $\tau_i \cap \tau_j \neq \emptyset$.

Кроме того, если $\tau_i \cap \tau_j \neq \emptyset$, то подграфы $G|_{\tau_j \cup \tau_i}$ и $G|_{\tau_j \cap \tau_i}$ также являются просто-вложенными в G . Таким образом, $\sigma_i \cap \sigma_j \in \widehat{FP}(G|_{\tau_j \cap \tau_i})$. Если при этом $\sigma_i \cap \sigma_j \neq \tau_i \cap \tau_j$, то $\sigma_i, \sigma_j \notin \widehat{FP}(G|_{\tau_j \cup \tau_i})$.

3. Если $\{\tau_1, \dots, \tau_N\}$ — просто вложенное покрытие графа G , и для каждого $i \in [N]$ выполнено $\sigma \cap \tau_i \in FP(G|_{\tau_i})$, то

$$\sigma \in FP(G) \Leftrightarrow \sigma \in FP(G|_{\sigma}).$$

Таким образом, эти шесть правил можно назвать базовыми инструментами для анализа фиксированных точек в TLN. Они позволяют:

1. Проверять, может ли подмножество нейронов быть поддержкой фиксированной точки.
2. Исключать невозможные поддержки (по сохранению или доминированию).
3. Собирать фиксированные точки из более простых подграфов.
4. Делать выводы о глобальной структуре сети на основе локальных свойств.

Таким образом, эти шесть элементарных правил образуют основу для анализа фиксированных точек в порогово-линейных нейронных сетях. С их помощью можно не только проверять, какие подмножества нейронов могут быть поддержками фиксированных точек, но и исключать невозможные варианты, собирать фиксированные точки по частям и делать выводы о поведении всей сети, глядя на её подграфы. Таким

образом, они дают простой и понятный способ изучения динамики TLN, опираясь только на структуру графа.

6. Правила склеивания

В предыдущей главе были рассмотрены способы определения возможных фиксированных точек и их поддержек на основании структуры сети, представленной в виде ориентированного графа. Однако если структура оказывается более сложной, этот анализ может становиться затруднительным. В таких случаях можно разбить граф на несколько частей и изучить их по отдельности, а затем восстановить фиксированные точки исходной сети. Для этого используются правила склеивания, которые мы рассмотрим в этом разделе.

Для того, чтобы установить связи между полным графом сети и его подграфами введем понятие просто-вложенного покрытия.

G – граф n вершин. Множество $U = \{\tau_1, \dots, \tau_N\}$ называется просто-вложенным покрытием G если выполняется:

1. Каждое τ_i – просто-вложено в G
2. Для каждой вершины $j \in [n]$ $\exists i \in [N]$ такое что $j \in \tau_i$, то есть U покрывает все вершины графа.[2]

Кроме того, можно ввести еще одно определение:

Множество $U = \{\tau_1, \dots, \tau_N\}$ называется просто-вложенным разбиением, если оно является просто-вложенным покрытием и τ_i попарно не пересекаются.[2]

Для каждого графа G можно найти тривиальное просто-вложенное покрытие, если взять $N = n$ и $\tau_i = \{i\}$ (что так же является просто-вложенным разбиением) или $N = 1$ и $\tau_i = [n]$. Однако такие покрытия не дают полезной информации о G , а более сложные, нетривиальные просто-вложенные покрытия, в свою очередь могут накладывать различные ограничения на множество фиксированных точек $FP(G)$. Некоторые из этих ограничений могут быть описаны правилами склеивания.

Нас интересует, при каких условиях локальные фиксированные точки $\sigma_i \in FP(G|_{\tau_i})$ можно склеить в глобальную фиксированную точку $\sigma = \cup \sigma_i \in FP(G)$. [2]

Для этого введем понятие комплекса склеивания:

$$F_G(U) = \{\sigma = \cup_i \sigma_i | \sigma \neq \emptyset, \sigma_i \in FP(G|_{\tau_i}) \cup \{\emptyset\}, \sigma_i \cap \tau_j = \sigma_j \cap \tau_i \forall i, j \in [N]\}$$

По сути это множество содержит все $\sigma \subseteq [n]$, которые можно получить, если согласованно склеивать локальные фиксированные точки подграфов. Согласованность означает, что на пересечениях множеств $\tau_i \cap \tau_j$ локальные поддержки σ_i и σ_j совпадают. Тогда можно переопределить $F_G(U)$:

$$F_G(U) = \{\sigma \neq \emptyset | \sigma \cap \tau_i \in \widehat{FP}(G|_{\tau_i}) \forall \tau_i \in U\}$$

Когда ввели это множество можем рассмотреть следующую лемму:

Лемма 4

Если G – граф CTLN, U – его просто-вложенное покрытие. Тогда:

$$FP(G) \subseteq F_G(U). [2]$$

Теперь определим более строгое множество в котором $\sigma \cap \tau_i \neq \emptyset$:

$$F_G^*(U) = \{\sigma \subseteq [n] | \sigma \cap \tau_i \in FP(G|_{\tau_i}) \forall \tau_i \in U\}$$

Это множество может оказаться пустым даже тогда, когда $F_G(U)$ содержит хотя бы один элемент. [2]

6.1. Связные объединения

Следующий важный вопрос состоит в том при каких условиях каждое согласованное семейство локальных фиксированных точек действительно соответствует глобальной фиксированной точке в G . Ответ на него даёт теорема 3, которая будет рассмотрена ниже, но сначала введем несколько определений.

Пусть $U = \{\tau_i\}_{i=1}^N$ – покрытие, тогда нерв покрытия – это множество

$$\mathcal{N}(U) = \{\alpha \subseteq [N] | \cap_{i \in \alpha} \tau_i \neq \emptyset\}$$

То есть нерв отражает, какие подмножества τ_i имеют общие элементы. [2]

Граф G называется связным объединением индуцированных подграфов $\{(G|_{\tau_i})\}$ если $\{\tau_1, \dots, \tau_N\}$ образуют просто-вложенное покрытие G , которое так же является связным(нerv связан). [2]

Перейдем непосредственно к теореме.

Теорема 3

Если граф G – связное объединение индуцированных подграфов $G|_{\tau_i}$, то есть U – связное и просто вложенное покрытие, то:

$$FP(G) = F_G(U). [2]$$

То есть, все возможные согласованные склейки локальных фиксированных точек являются фиксированными точками всей сети.

Эту теорему можно объяснить тем, что если граф G имеет связное просто-вложенное покрытие, то мы можем пронумеровать множества τ_1, \dots, τ_N так, что благодаря связности выполняется условие $\tau_l \cap (\tau_1 \cup \dots \cup \tau_{l-1}) \neq \emptyset \forall l$. Тогда по лемме о пересечениях и объединениях(см Лемма 2, 5.4) каждое из объединений $\tau_1 \cup \tau_2 \cup \dots \cup \tau_l$ тоже будет просто-вложенным в G . Это позволяет применять попарное склеивание шаг за шагом: склеить фиксированные точки на первых двух подграфах τ_1 и τ_2 , затем результат склеить с τ_3 и так далее. Таким образом, мы получаем фиксированную точку на всём объединении. [2]

Пример 1

Рассмотрим следующий пример для графа G , изображенного на Рис. 4. чтобы показать

действие теоремы.

$U = \{\tau_1 = \{1,2,3\}, \tau_2 = \{3,4,5\}, \tau_3 = \{5,6,4,8\}\}$ – его просто-вложенное связное покрытие, которое можно рассмотреть на Рис.4А. Заметим, что каждое τ_i определяет индуцированный подграф $G|_{\tau_i}$ для которых легко найти фиксированные точки. Для этого воспользуемся уже изученными правилами.

1. Подграф $G|_{\tau_1}$ – цикл из трех вершин и по правилу 3 о циклах фиксированная точка – это $\{1,2,3\}$
2. $G|_{\tau_2}$ состоит из клики $\{3,4\}$ и независимой точки $\{5\}$. Тогда по графовому правилу 2 о кликах $\{3,4\} \in FP(G|_{\tau_2})$, $\{5\}$ – независимое множество и тривиальный сток, так что является фиксированной точкой по правилу 1. Тогда по правилу 6 о стоке вне поддержки $\{3,4,5\} \in FP(G|_{\tau_2})$ так как уже знаем что $\{3,4\} \in FP(G|_{\tau_2})$.
3. Подграф $G|_{\tau_3}$ состоит из нескольких циклов и по правилу 3 о циклах $\{5,6,7\}, \{6,7,8\}, \{5,6,7,8\} \in FP(G|_{\tau_3})$.

Склеивая совместимые поддержки на пересечениях (например, $3 \in \tau_1 \cap \tau_2$, $5 \in \tau_2 \cap \tau_3$), можно получить глобальные фиксированные точки сети:

$$FP(G) = F_G(U) = \{1234, 1234678, 123456, 12345678, 567, 5678, 678\}$$

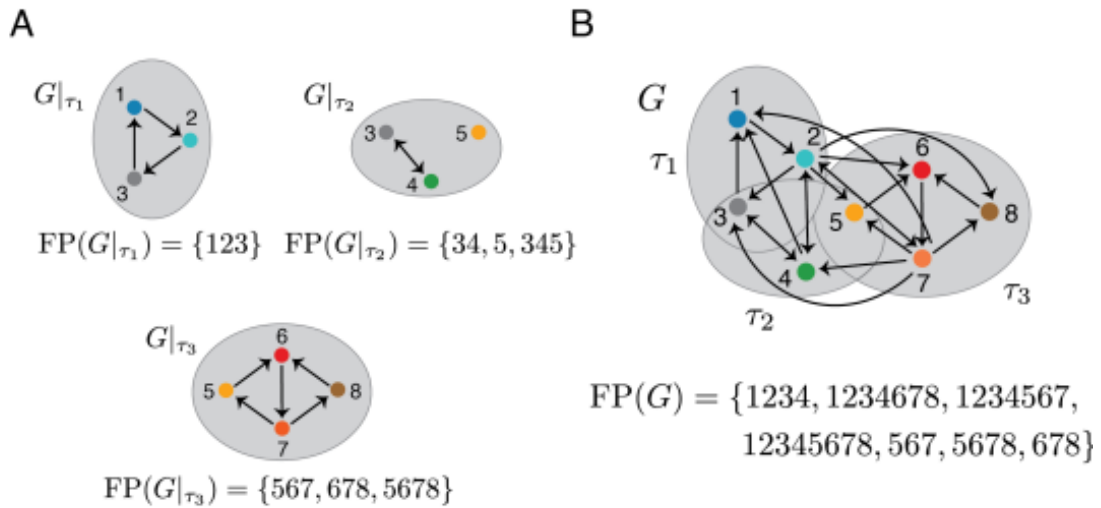


Рис.4. Пример склейки фиксированных точек [2]

А. Подграфы $G|_{\tau_i}$ индуцированные просто-вложенными связными покрытиями графа G и их фиксированные точки. В. Граф G и множество его фиксированных точек.

6.2. Просто-вложенные разбиения.

В предыдущем разделе были рассмотрены случаи связных покрытий. Однако может возникнуть иной случай – просто-вложенные разбиения. Они состоят из попарно непересекающихся подмножеств τ_i . Рассмотрим некоторые конструкции, которые возникают в результате просто-вложенных разбиений.

Пусть граф G состоит из подграфов $G|_{\tau_i}$, соответствующих разбиению $U = \{\tau_1, \dots, \tau_N\}$. Тогда можно рассмотреть несколько случаев их примеры представлены на Рис.5:

1. G – непересекающееся объединение, если между компонентами нет ребер.
2. G – объединение-клика, если между любыми τ_i и τ_j таких что $i \neq j$ существуют все возможные ребра.
3. G – циклическое объединение, если он содержит все возможные ребра из τ_i в τ_{i+1} для $\forall i \in \{1, \dots, N-1\}$, а так же ребра из τ_N в τ_1 .
4. G – линейная цепь, если в нем есть все возможные ребра из τ_i в τ_{i+1} для $\forall i \in \{1, \dots, N-1\}$ и между различными τ_i и τ_j нет других ребер. [2]

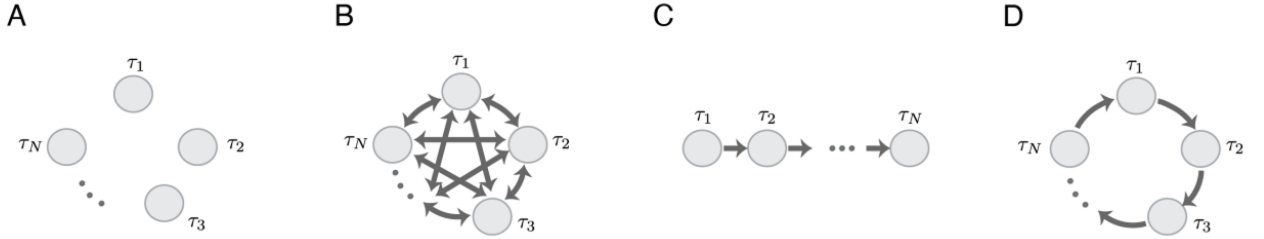


Рис.5. Примеры непересекающегося объединения, объединения-клики, линейной цепи и циклического объединения. [2]

Каждая из этих архитектур позволяет описывать множество фиксированных точек всего графа G на основе фиксированных точек его подграфов. Особенно просты в этом плане непересекающиеся объединения, так как в них отсутствуют рёбра между подграфами, и каждая компонента функционирует независимо. Значит, любые сочетания фиксированных точек подграфов дают фиксированную точку всего графа. Это утверждение сформулировано в следующей теореме.

Теорема 4

Если G – непересекающееся объединение подграфов $G|_{\tau_i}$, которые соответствуют вложенному разбиению $U = \{\tau_1, \dots, \tau_N\}$. Тогда :

$$FP(G) = F_G(U) = \{\cup \sigma_i | \sigma_i \in \widehat{FP}(G|_{\tau_i}) \forall i\} \setminus \{\emptyset\}$$

Таким образом, множество фиксированных точек G включает все непустые объединения фиксированных точек на компонентах, в том числе случаи, когда на некоторых компонентах активных нейронов нет (что соответствует $\sigma_i = \emptyset$). [2,4]

Кроме того, стоит отметить, что в этом случае множество фиксированных точек максимально возможное среди всех разбиений на просто-вложенные подграфы и выражается следующим образом:

$$|FP(G)| = \prod_{i=1}^N (|FP(G|_{\tau_i})| + 1) - 1$$

Но не во всех случаях допускаются пустые подмножества σ_i : например в объединении-клике или циклическом объединении, так как в таких графах обязательно должны быть ребра между компонентами. Это обобщено в еще одной теореме.

Теорема 5

Если граф G – объединение-клика или циклическое объединение подграфов $G|_{\tau_i}$, которые соответствуют вложенному разбиению $U = \{\tau_1, \dots, \tau_N\}$. Тогда :

$$FP(G) = F_G^*(U) = \{\cup \sigma_i | \sigma_i \in FP(G|_{\tau_i}), \forall i \in [N]\}$$

Таким образом, допустимы только такие склейки, в которых каждый подграф вносит ненулевой вклад, а общее количество фиксированных точек определяется произведением мощностей множеств фиксированных точек подграфов[2,4]:

$$|FP(G)| = \prod_{i=1}^N |FP(G|_{\tau_i})|$$

Не меньший интерес представляют линейные цепи. В отличие от предыдущих случаев здесь важен порядок склейки, так как по правилу 5(ii) каждая σ_i не сохраняется в $FP(G|_{\tau_i \cup \tau_{i+1}})$. В свою очередь σ_{i+1} сохраняется при подобной операции. В результате поэтапного склеивания получим следующую картину: $FP(G|_{\tau_1 \cup \dots \cup \tau_l}) = FP(G|_{\tau_l})$. Это приводит нас к новой теореме.

Теорема 6

Если G – линейная цепь подграфов $G|_{\tau_i}$, которые соответствуют вложенному разбиению $U = \{\tau_1, \dots, \tau_N\}$. Тогда :

$$FP(G) = FP(G|_{\tau_N})$$

И, следовательно, $|FP(G)| = |FP(G|_{\tau_N})|$.

Таким образом, структура линейной цепи резко ограничивает фиксированное поведение: сохраняются только фиксированные точки последнего компонента.[2,4]

Итак, в зависимости от типа разбиения графа множество $FP(G)$ может существенно различаться: от полного набора сочетаний фиксированных точек подграфов до случаев, когда сохраняется лишь последний элемент разбиения.

6.3. Основные мотивы (core motifs)

Если граф допускает разбиение на подграфы известной структуры, это позволяет исследовать, какие из них являются основными мотивами. Вспомним, что основной мотив – это подграф, поддерживающий ровно одну фиксированную точку с полной поддержкой: $FP(G|_{\sigma}) = \{\sigma\}$.

Обозначим множество всех основных мотивов графа G которые остаются после склеивания как:

$$FP_{core}(G) = \{\sigma \in FP(G) | G|_{\sigma} - \text{основной мотив } G\}$$

Рассмотрим несколько следствий о основных мотивах, которые вытекают из уже рассмотренных теорем.

Следствие 2

Если граф G – связное объединение подграфов, и каждая компонента $G|_{\tau_i}$ является основным, тогда G также является основным мотивом.[2]

Это следствие опирается на теорему 4, описывающую структуру фиксированных точек в случае непересекающегося объединения.

Следствие 3

Если граф G объединение-клика или циклическое объединение компонент τ_1, \dots, τ_N . Тогда множество основных мотивов определяется как:

$$FP_{core}(G) = \{\cup \sigma_i | \sigma_i \in FP_{core}(G|_{\tau_i})\}$$

В частности, граф G сам является основным мотивом тогда и только тогда, когда каждая компонента $G|_{\tau_i}$ является основным мотивом.[2]

Это следствие основывается на теореме 5, где рассматриваются кликовые и циклические объединения. Согласно теореме, фиксированные точки всего графа получаются путём объединения фиксированных точек его компонент, при условии, что каждая из них непустая. Если каждая компонента является основным мотивом и имеет единственную фиксированную точку, равную всему своему множеству, то объединение этих точек дает только множество всех вершин. Это означает, что и весь граф в таком объединении является основным мотивом.

Таким образом, большие основные мотивы могут быть построены из меньших, используя описанные выше правила склейки. Каждый из таких мотивов предсказывает определённый тип аттракторов сети.

Рассмотрим как все это может быть использовано. Такие архитектуры применяются для моделирования явлений с последовательными аттракторами, например:

1. Хранения последовательной информации и ее воспроизведения
2. Центральных генераторов паттернов(CPG). Это нейронные сети, которые могут генерировать активности с ритмичными паттернами, такими как бег, ходьба, дыхание, полет птицы и так далее.

Рассмотрим пример такой сети на Рис.6.

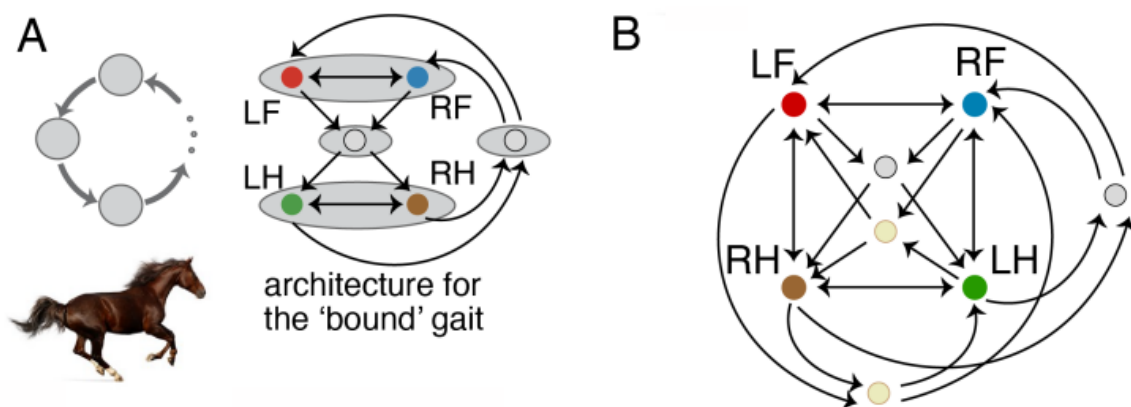


Рис.6. Пример CPG.

А. Архитектура скачка. В. Архитектура сети в которая допускает скачки и рысь.

Сеть изображенная на этой схеме моделирует бег четвероногого животного. На Рис. 6А представлен граф на 6 вершинах, моделирующий паттерн скачка. Каждая вершина ассоциирована с одной ногой животного. Клики между парными ногами (то есть клика

между парой вершин соответствующих передним ногам и клика между парой вершин, соответствующих задним) обеспечивают одновременную активацию, а циклическое объединения гарантирует последовательность активности.

Аналогичную структуру можно построить для моделирования паттерна рыси. В этом случае мы опять получаем граф на 6 вершинах, в котором кликами соединены противоположные ноги, то есть правая передняя – с левой задней и левая передняя – с правой задней. Изобразим соответствующую сеть на Рис.7.

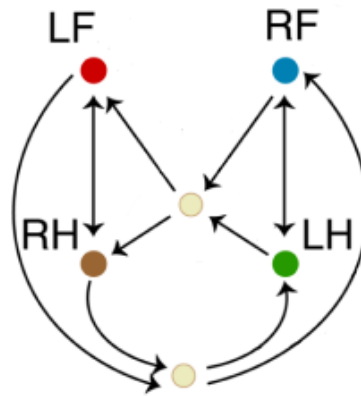


Рис.7. Архитектура бега рысью

На Рис. 6В показана комбинированная сеть, которая объединяет паттерны рыси и скачка. Такая конструкция создаётся путём наложения двух отдельных архитектур. Несмотря на то, что в объединённой сети подграфы для каждого паттерна уже не являются чистыми циклическими объединениями (так как добавлены рёбра), сеть всё равно воспроизводит ограниченные циклы, аналогичные тем, что возникают в исходных подграфах.

Это приводит к важному наблюдению: даже после наложения и усложнения структуры, динамика может сохранять черты базовых циклических объединений.

7. Заключение

В рамках данной работы было проведено исследование динамики рекуррентных нейронных сетей (RNN), с акцентом на комбинаторные порогово-линейные нейронные сети (CTLNs). Основное внимание было уделено исследованию динамики таких сетей через анализ фиксированных точек и аттракторов, возникающих в зависимости от топологии графа.

Важной частью исследования стали графовые правила, которые позволяют понять динамику сети, глядя только на её структуру. Были рассмотрены графы с равномерной входящей степенью, случаи графового доминирования и простые подграфы, которые упрощают анализ. Эти правила позволяют предсказывать наличие или отсутствие фиксированных точек только по структуре графа без необходимости рассматривать иные параметры

Так как некоторые графы могут быть слишком сложными, также были изучены правила склеивания – методы объединения подграфов для анализа больших сетей. Показано, что такие подходы позволяют собирать глобальную динамику сети из более простых компонентов.

Таким образом, в ходе работы была достигнута поставленная цель — изучить зависимость динамики рекуррентных нейронных сетей от структуры графа и правил склеивания. Работа позволила значительно углубить понимание связи между архитектурой сети и её динамическими свойствами и дала более широкое представление о том, как локальные свойства подграфов могут определять глобальное поведение всей сети.

8. Список литературы

1. Christopher Olah, “Understanding LSTM Networks”, colah's blog, 2015
2. Carina Curto, Katherine Morrison, “Graph rules for recurrent neural network dynamics: extended version”, 2023
3. By Kathryn Ippolito, “Balancing Act in the Brain: Excitatory and Inhibitory Activity”, 2020
4. Carina Curto, Jesse Geneson, Katherine Morrison, “Stable fixed points of combinatorial threshold-linear networks”, November 18, 2023
5. Daniela Egas Santander, Stefania Ebli, Alice Patania, Nicole Sanderson, Felicia Burtscher, Katherine Morrison, Carina Curto, “Nerve Theorems for Fixed Points of Neural Networks”, 2022