

## Спайковые НС

### Spike Response Model

**Общее описание.** Spike Response Model (SRM) - наиболее популярная модель спайкового нейрона. SRM своей популярностью обязана простотой математической интерпретации - вся динамика нейрона описывается одним уравнением вида  $u(t)$ , которое описывает напряжение на мембране нейрона и, по сути, является решением дифференциального уравнения для моделей типа *Integrate and fire*. Динамику моделей *Integrate and fire* можно описать так: нейрон суммирует входные сигналы и по достижению определенного порога, производит спайк, после чего нейрон переходит в состояние рефракторности, находясь в котором, вероятность нового спайка крайне мала.

Поведение описанное выше можно поэтапно собрать в одну формулу:

1. *Функция описывающая напряжение на синапсах.* В качестве такой функции можно взять альфа функцию с экспоненциальным подъёмом и спадом, причем подъем и спад наиболее натуральным будет взять быстрым и медленным соответственно. Типичный график подобной функции можно посмотреть на рисунке ниже:

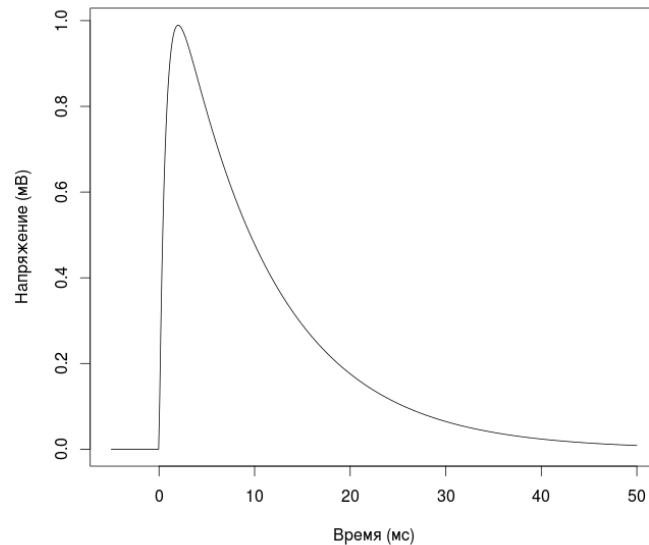


Рис. 1: Потенциал на синапсах

Такая функция задается формулой:

$$\epsilon(t) = \epsilon_0(\exp(-t/t_m) - \exp(-t/t_s)), \quad (1)$$

где  $\epsilon_0 = 1.3\text{мВ}$  - константа задающая масштаб потенциала,  $t_m = 0.7\text{мс}$  - константа отвечающая подъём,  $t_s = 10\text{мс}$  - константа отвечающая спад.

2. *Функция описывающая рефракторность нейрона.* Основное требования к такой функции в том, чтобы напряжение на нейроне резко падало вниз, и потом медленно восстанавливалось. График подобной функции можно увидеть на рисунке ниже:

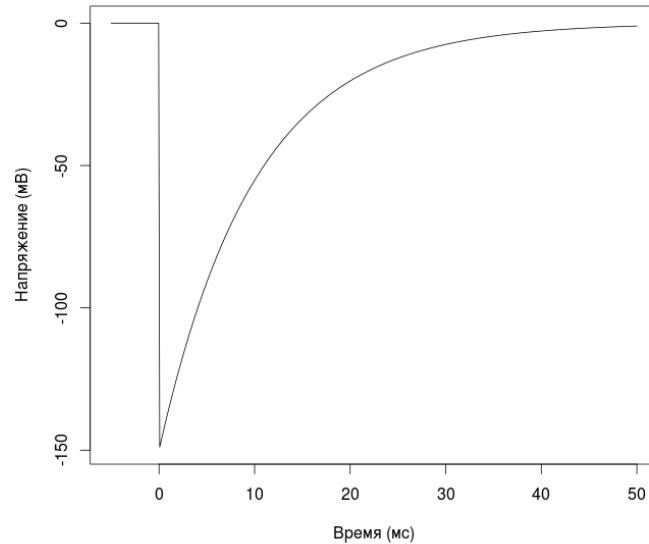


Рис. 2: Рефракторность нейрона

Здесь используется данная функция:

$$\eta(t) = \eta_0(-\exp(-t/t_m)), \quad (2)$$

где  $\eta_0 = -150\text{мВ}$  - константа описывающая минимальное напряжение на мембране, от которого идёт медленное восстановление,  $t_m$  - скорость восстановления можно взять из функции синаптического потенциала, для простоты.

В итоге, используя эти функции, можно записать уравнение, которое будет описывать напряжение на мембране нейрона, причём:

- Пусть нейрон  $i$  имеет  $N$  синапсов и у каждого синапса есть вес  $w_i$ , тогда напряжение на мембране в данный момент времени  $t$  будет взвешенной суммой синаптических потенциалов:  $\sum_{j=1}^N w_j \sum_{f_j} \epsilon_j(t - f_j)$ , где  $f_j$  - время спайка на синапсе  $j$ .
- Рефракторность нейрона будет простой суммой по всем спайкам, которые произвел нейрон  $i$ :  $\sum_{f_i} \eta(t - f_i)$
- Нейрон имеет т.н. потенциал покоя. Биологические нейроны имеют разнообразные значения такого потенциала, как правило берут  $u_{rest} = 70 \text{ мВ}$ .

Таким образом получаем формулу, которая объединяет все вышеописанные особенности:

$$u(t) = u_{rest} + \sum_{j=1}^N w_j \sum_{f_j} \epsilon_j(t - f_j) + \sum_{f_i} \eta(t - f_i), \quad (3)$$

Типичный график иллюстрирующий работу функции 3 ниже:

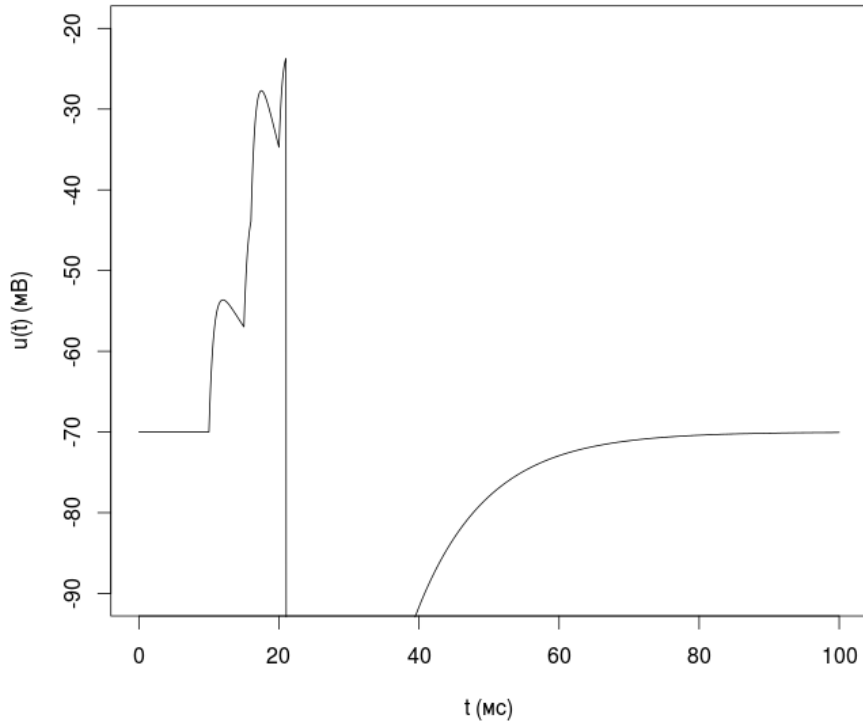


Рис. 3: Потенциал нейрона

На рисунке представлен случай когда произошли спайки на двух синапсах во времена  $f_1 = \{10, 16\}$  и  $f_2 = \{15, 20\}$ , которые заставили нейрон произвести спайк в  $f_i = \{21\}$ . Веса были выбраны большими, для наглядности графика.

**Генерация спайков.** Модель описанная выше включает в себя только поведение при данных временах спайков на синапсах нейрона и самого нейрона. Рассмотрим вопрос условия генерации спайков.

Классическим вариантом модели генерации спайков является модель с порогом напряжения. Т.е. при достижении нейроном какого-то конкретного порогового напряжения - нейрон генерирует спайк. Как показала практика, наиболее удобным в математическом анализе таких моделей является

стохастическая модель порога. Особенность такой модели в том, что нейрон генерирует спайк с определенной вероятностью, которая всё же зависит от напряжения на мембране и которая делает резкий скачок около порогового напряжения. Пример такой зависимости на рисунке ниже:

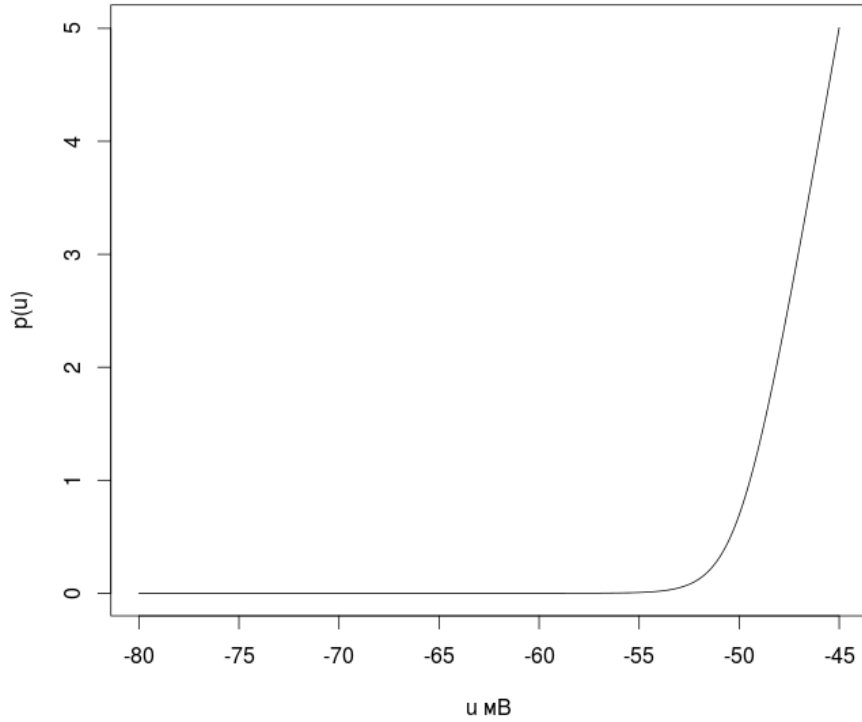


Рис. 4: Плотность вероятности генерации спайка

Такая плотность вероятности является ненормализованной, т.к. её интеграл больше единицы, и имеет все свойства плотности вероятности Пуассоновского процесса. Например, вероятность генерации одного спайка в отрезок времени  $\Delta t$  будет  $P = p(u(t))\Delta t$ . В данной работе взята функция вида

$$p(u) = \frac{\beta}{\alpha} \ln[1 + \exp(\alpha(U_{thresh} - v))] - \alpha(U_{thresh} - v) \quad (4)$$

где константы  $\alpha = \beta = 1$ , и порог  $U_{thresh}$  по достижению которого функция резко растёт вверх равен -50 мВ.

**Математический фреймворк.** Используя выкладки выше можно вывести математическую базу для Spike Response Model. Ниже приведены основные формулы.

- Вероятность отсутствия спайков в диапазон  $[0, T]$  при данных

спайках на синапсах  $X$  ( $Y_0 = \{\}$ ):

$$P(Y_0|X) = S[0, T] = \int_0^T \exp(-p(t)) dt, \quad (5)$$

- **Вероятность генерации спайковой последовательности  $Y = \{t_{f1}, t_{f2}, \dots, t_{fn}\}$  диапазоне  $[0, T]$  при данных спайках на синапсах  $X$ .**

$$P(Y|X) = S[0, t_{f1}] p(t_{f1}) S[t_{f1}, t_{f2}] p(t_{f2}) \dots p(t_{fn}) S[t_{fn}, T], \quad (6)$$

что можно упростить, и считать только один интеграл:

$$P(Y|X) = S[0, T] p(t_{f1}) p(t_{f2}) \dots p(t_{fn}) \quad (7)$$

**Обучение.** При соединении таких SRM нейронов в сеть, имея набор входных спайков для каждого нейрона  $X_1 = \{t_{f1}^1, t_{f2}^1, \dots, t_{fn}^1\}$ ,  $X_2 = \{t_{f1}^2, t_{f2}^2, \dots, t_{fn}^2\}$  и так далее, сеть будет порождать в определенной степени уникальный ответ в виде выходных спайков  $Y = \{t_{f1}^Y, t_{f2}^Y, \dots, t_{fn}^Y\}$ . Главный вопрос, который нужно поставить перед тем как обучать нашу модель - вопрос каким должен быть выходной ответ сети, и как он должен измениться во время обучения.

Поставим задачу классификации и зададим требование, чтобы ответ сети для одного класса входных данных отличался от другого класса. Для выполнения этого требования необходимо, чтобы сеть выдавала наиболее постоянный, но, разумеется, не уникальный ответ, на все виды входных данных. Сеть построенная на SRM обладает не малой долей стохастичности и как одна из главных задач обучения таких сетей является задача снижения этой стохастичности, или иначе называя энтропии нейрона.

Энтропию нейрона, как меру непредсказуемости его поведения, можно вывести используя классические уравнения из Теории Информации. Пусть  $\xi$  - возможный ответ нейрона вида  $[\sigma(0), \sigma(\Delta t), \dots, \sigma(T)]$ , где  $\sigma(t) = 0$  если в данный момент нейрон не произвел спайк и  $\sigma(t) = 1$ , если спайк был. Множество всех возможных вариантов спайковых последовательностей обозначим через  $\Omega$ , причем  $\xi \in \Omega$ . Условная энтропия, при данном входе  $X$ , будет равна

$$H(\Omega|X) = - \int_{\Omega} p(\xi) \log(\xi) d\xi \quad (8)$$

где  $p(\xi)$  - плотность вероятности спайковых последовательностей  $\xi$ .

Снижая энтропию нейрона, нейрон обучается генерировать строго определенный набор выходных спайков для определенного входа, и в зависимости от разнородности входных спайковых последовательностей, нейрон будет генерировать разные выходные последовательности для разных классов. Перед тем как попытаться снизить энтропию нейрона, её необходимо измерить.

Измерение энтропии всегда было трудоемкой задачей, так как влечет за собой измерение полной плотности вероятности. Математический фреймворк описанный выше позволяет вывести вероятность спайковой последовательности как функцию от входа  $X$  и весов нейрона. Из-за рефракторности, нейрон, может производить ограниченный набор спайков в определенный момент времени и при сужении этого интервала, вероятность большого

числа спайков становится меньше. Эту особенность можно использовать в контексте численного интегрирования и посчитать интеграл только для ограниченного набора  $\xi$ :

- При отсутствии спайков на  $[0, T]$

$$H_0(\Omega_0|X) = -p_0 \log(p_0) \quad (9)$$

где  $p_0 = S[0, T]$ , смотри 7.

- При наличии одного спайка на  $[0, T]$

$$H_1(\Omega_1|X) = - \int_0^T p_1(t_{f1}) \log(p_1(t_{f1})) dt_{f1} \quad (10)$$

где  $p_1(t_{f1}) = S[0, t_{f1}] p(t_{f1}) S[t_{f1}, T]$

- При наличии двух спайков на  $[0, T]$

$$H_2(\Omega_2|X) = - \int_0^T \int_{t_{f2}}^T p_2(t_{f1}, t_{f2}) \log(p_2(t_{f1}, t_{f2})) dt_{f1} dt_{f2} \quad (11)$$

где  $p_2(t_{f1}, t_{f2}) = S[0, t_{f1}] p(t_{f1}) S[t_{f1}, t_{f2}] p(t_{f2}) S[t_{f2}, T]$

В итоге при выборе  $T$ , например, 20 мс, подсчет интеграла будет сводиться к сумме

$$H(\Omega|X) = H_0 + H_1 + H_2 \quad (12)$$

причем энтропия нейрона при наличии трех спайков и более стремится к нулю, так как их вероятность предельно мала.

**Минимизация энтропии.** Функция вида позволяет посчитать градиент изменения энтропии нейрона  $i$  как функции весов

$$\frac{dH(\Omega|X)}{dw_{ij}} = - \int_{\Omega} p(\xi) (\log(\xi) + 1) \frac{d\log(p(\xi))}{dw_{ij}} d\xi \quad (13)$$

причем

$$\frac{d\log(p(\xi))}{dw_{ij}} = - \int_0^T p'_i(t) \sum_{f_j \in \xi} \epsilon(t|f_j) dt + \sum_{f_i \in Y} \frac{p'_i(f_i)}{p_i(f_i)} \epsilon(f_i|f_j) \quad (14)$$

где

$$p'_i(t) = \frac{dp_i(t)}{du_i(t)} = \frac{\beta}{1 + \exp(\alpha(U_{thresh} - u_i(t)))} \quad (15)$$