

# НЕЙРОСЕТЕВОЕ МОДЕЛИРОВАНИЕ ПРОЦЕССОВ ПЕРЕНОСА ЭНЕРГИИ В СИСТЕМАХ МЕМБРАННЫХ ПРОТЕИНОВ

**П. В. Назаров**

## **Введение**

В последнее время в литературе появилось много работ, посвященных методу определения неизвестных параметров различных систем с помощью имитационного моделирования, в частности моделирования методом Монте-Карло. Применение такого подхода имеет смысл в том случае, когда поведение системы трудно либо невозможно описать аналитически, но имеется информация об элементарных процессах, происходящих в ней.

Однако при этом возникает ряд трудностей, наиболее значимыми из которых являются большие вычислительные затраты при практической реализации имитационного моделирования. Особенно остро это проявляется при попытке использования имитационной модели в методах оптимизации параметров, когда число пусков имитационного моделирования равняется числу вычислений функции невязок.

Нами предложен метод замены имитационной модели системы искусственной нейронной сетью (ИНС) для ускорения алгоритмов оптимизации. Разработанный метод был применен для анализа процессов переноса энергии в таких сложных бимолекулярных системах, как системы мембранных протеинов.

## **1. Теория**

Рассмотрим следующую задачу. У некоторой экспериментальной системы имеется набор входных параметров  $P$ , часть из которых из-

вестна ( $P_0$ ), а часть предстоит определить ( $P_X$ ), и набор выходных значений  $F$ . В этом случае можно сказать, что рассматриваемая система выполняет некоторое преобразование  $\Theta$ :

$$\Theta(P_0, P_X) = F. \quad (1)$$

Пусть для этой системы можно построить адекватную имитационную модель. Тогда эта модель будет выполнять преобразование  $\Xi$ :

$$\Xi(P_0, P_X) = F^*. \quad (2)$$

Алгоритм определения параметров с использованием имитационного моделирования выглядит следующим образом:

1. Экспериментально получают некоторую выборку выходных значений  $F$  при различных входных параметрах системы.
2. Делаются начальные приближения неизвестных параметров  $P_X$ .
3. Запускается алгоритм оптимизации, который, изменяя значения  $P_X$ , минимизирует ошибку.

Как уже упоминалось, наиболее серьезной проблемой в этой схеме являются временные затраты на выполнение имитационного моделирования. В некоторых случаях это приводит к тому, что подобный подход неприменим вообще из-за слишком большого (несколько месяцев) времени выполнения алгоритма оптимизации. Для того чтобы существенно ускорить этот процесс, нами было предложено заменить модель искусственной нейронной сетью.

Из литературы [1, 2] известно, что непрерывные функции могут быть с любой, наперед заданной точностью аппроксимированы линейной комбинацией и суперпозицией сигмоидальных функций, т. е. многослойной ИНС. В нашем случае это означает, что операция имитационного моделирования  $\Xi$  может быть аппроксимирована нейросетевым преобразованием

$$\Psi(P_0, P_X) = F^*. \quad (3)$$

Вычислительные затраты при этом будут значительно ниже, чем при вычислении результатов имитационного моделирования.

Предлагаемый подход представлен на рис. 1.

Перед началом работы необходимо, используя имитационную модель, создать обучающую выборку и обучить ИНС. Причем и генерацию обучающей выборки, и обучение приходится повторять при любом изменении модели, например при ее коррекции либо усложнении.

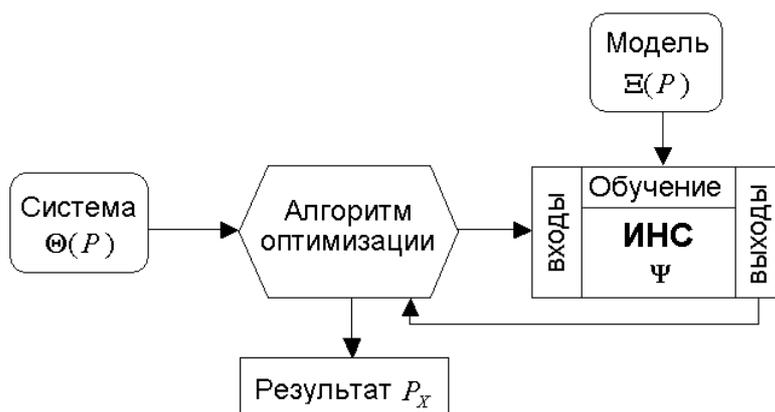


Рис.1. Применение ИНС для замены имитационной модели

### 3. Описание модели

Описанный метод был использован нами для моделирования и анализа процессов диполь-дипольного переноса энергии между флуоресцентными метками мембранных протеинов.

При построении формализованной модели липидная мембрана аппроксимировалась структурой, приведенной на рис. 2а. Каждый протеин заменялся моделью, представленной на рис. 2б.

В биоорганической химии для определения расстояний внутри и между различными макромолекулами часто используется метод так называемой спектроскопии резонансного переноса энергии [3]. Суть метода заключается в следующем. В исследуемые макромолекулы (в данном случае – мембранные протеины) внедряют флуоресцентные зонды двух типов: доноры (D) со сравнительно большим временем жизни в возбужденном состоянии и акцепторы (A). Доноры возбуждаются внешним источником освещения. Часть из них передает свою энергию акцепторам, которые ее излучают с большей длиной волны.

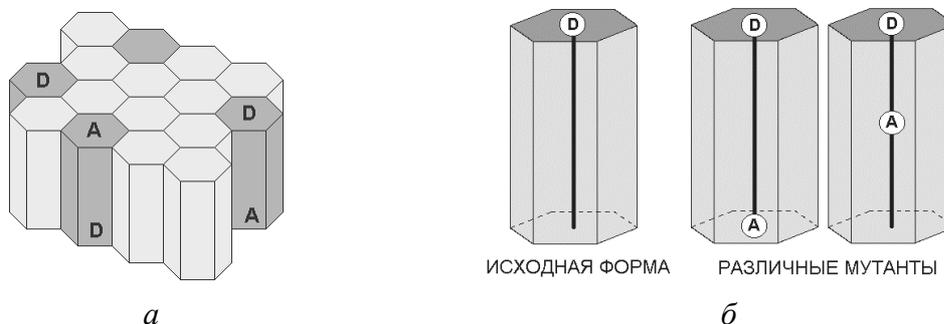


Рис 2. Модель мембраны (а) и мембранных протеинов (б) с флуоресцентными зондами

Вероятность переноса энергии в системе донор-акцептор обратно пропорциональна 6-й степени расстояния между ними.

#### 4. Результаты применения ИНС

Для замены имитационной модели мы использовали трехслойный перцептрон. Необходимое число нейронов в слоях устанавливалось экспериментальным путем. ИНС обучалась методом обратного распространения ошибки. Для избежания переобучения каждые 10 итераций работа сети тестировалась на контрольном множестве (200 элементов). Если ошибка при этом не уменьшалась в течение некоторого времени, обучение останавливалось.

В качестве примера нейросетевой аппроксимации на рис. 3 представлена зависимость эффективности переноса энергии от положения акцептора в липидном слое. Кружками обозначены результаты имитационного моделирования, линиями – результаты работы сети. Погрешность аппроксимации составила порядка 2 %.

Следует отметить, что применение нейросетевого моделирования позволило ускорить получение результата в  $10^4$ – $10^5$ . Результатом работы ИНС является гладкая функция, что позволяет использовать градиентные методы оптимизации.

К недостаткам метода можно отнести то, что количество варьируемых параметров ограничено. Максимальное число параметров, при котором метод работал на нашей задаче, равнялось пяти. Это ограничение связано с проблемами создания репрезентативной обучающей выборки. Другой недостаток состоит в том, что аппроксимация всегда может содержать какие-либо погрешности.

Тем не менее метод всегда может быть использован для быстрого получения очень хороших начальных приближений при оптимизации и изучении функциональных зависимостей выходов системы от ее входов.

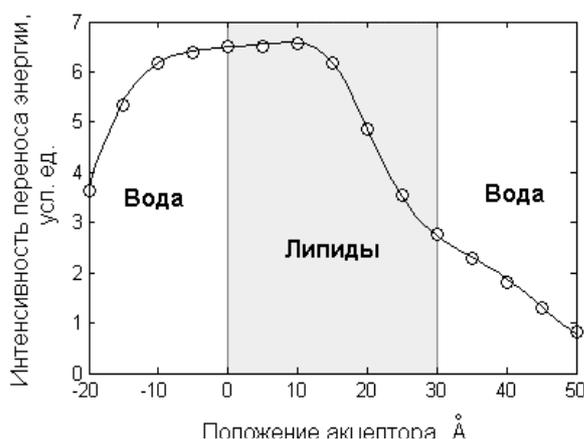


Рис. 3. Зависимость переноса энергии от положения флуоресцентного зонда. Кружками показана интенсивность переноса, рассчитанная с помощью имитационной модели, линией – нейросетевая аппроксимация

## Литература

1. *Cybenko G.* Approximations by superpositions of a sigmoidal function // *Math Contr Signals Syst.* 1989. V. 2. P. 304–314.
2. *Hornik K., Stinchcombe M., and White H.* Multilayer feedforward networks are universal approximators // *Neural Networks.* 1989. V. 2. P. 359–366.
3. *Koppel D., et al.* Intramembrane Position of Membrane-Bound Chromophores Determined by Excitation Energy Transfer // *Biochemistry.* 1979 V. 18. P. 5450–5457.