

КОМПЛЕКСНЫЙ АНАЛИЗ БОЛЬШИХ ДАННЫХ ПРИ ИССЛЕДОВАНИИ СЛОЖНЫХ БИОМОЛЕКУЛЯРНЫХ СОЕДИНЕНИЙ В СИСТЕМАХ ПРИКЛАДНОЙ ФЛУОРЕСЦЕНТНОЙ СПЕКТРОСКОПИИ

Н. Н. Яцков, В. В. Апанасович

Белорусский государственный университет, Минск, Беларусь

E-mail: yatskou@bsu.by

Разработан комплексный подход на основе методов имитационного моделирования и интеллектуального анализа больших данных для исследования сложных биомолекулярных соединений в системах прикладной флуоресцентной спектроскопии. Получены точные оценки времен затухания флуоресценции флуорофоров при обработке кинетических кривых затухания флуоресценции сложных молекулярных систем – ошибка снижена в два раза в сравнении с классическим анализом. Повышена эффективность анализа экспериментальных данных флуоресценции – выделены белковые комплексы, что трудно выполнимо традиционным анализом.

Ключевые слова: *флуоресцентная спектроскопия, комплексный подход, имитационное моделирование, интеллектуальный анализ данных.*

ВВЕДЕНИЕ

Методы экспериментальной флуоресцентной спектроскопии используются для изучения оптических свойств молекулярных соединений и находят широкое применение при исследовании фотонных искусственных материалов для приёма и передачи энергии электронного возбуждения, белковых комплексов, биополимеров, секвенировании молекул ДНК и РНК, определении биофункций генов и сетей их взаимодействия, биологических мембран, биомаркеров клеток и изменений в органических тканях [1-3]. Одно из основных ограничений, при обработке больших данных флуоресцентной спектроскопии, состоит в отсутствии универсальных эффективных автоматизированных систем поддержки и принятия решений, включающих статистические протоколы планирования и проведения экспериментов, программные средства обработки, интерпретации результатов и моделирования изучаемых биофизических процессов [4, 5]. Разработка новых усовершенствованных систем поддержки и принятия решений должна упростить и автоматизировать обработку экспериментальных измерений, повысить точность оценивания искомых параметров, расширить пределы интерпретации и предсказания физических процессов.

Целью работы является разработка комплексного подхода к анализу больших данных в системах обеспечения прикладной флуоресцентной

спектроскопии, включающего физические модели, методы и алгоритмы интеллектуального анализа данных (ИАД) [6, 7], программные средства для автоматизации процессов исследования молекулярных и клеточных систем [8].

КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ И ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ БОЛЬШИХ ДАННЫХ ПРИ ИССЛЕДОВАНИИ СЛОЖНЫХ МОЛЕКУЛЯРНЫХ СОЕДИНЕНИЙ И КЛЕТОК

Идея подхода к анализу больших данных состоит в изучении объекта исследования с применением имитационного моделирования биологических и физических процессов, протекающих в объекте исследования; сравнении смоделированных и наиболее информативных экспериментальных данных, отобранных методами по снижению размерности данных; определении параметров физических процессов с использованием алгоритмов оптимизации.

Предполагается, что через серию экспериментов, порождающих наборы больших данных, исследуется биофизический объект, сущностные признаки, или характеристики, которого, такие как набор биофизических параметров A , должны быть определены в ходе анализа данных. В комплексном подходе к наборам многомерных данных применяются алгоритмы ИАД с целью выделения наиболее информативных или существенных данных для последующего глубокого изучения и нахождения оценок параметров A с использованием имитационных моделей.

Схема исследования объекта с использованием Q экспериментов, представляющих некоторый глобальный эксперимент, с использованием комплексного подхода расположена на рисунке 1.

ПРИМЕНЕНИЕ АЛГОРИТМОВ КОМПЛЕКСНОГО ПОДХОДА

Проведена апробация алгоритмов комплексного подхода. Произведена оценка оптических параметров имитационно смоделированных наборов данных микроскопии визуализации времени жизни флуоресценции, представляющих три системы флуорофоров и характеризующихся одно-, двух- и стрэтчэкспоненциальными законами испускания флуоресценции. Применение разработанных алгоритмов позволило быстрее и более точно определить параметры биофизических и оптических процессов в молекулярных соединениях в сравнении с классическим способом анализа, основанным на отдельной обработке каждого набора данных. Ошибка оценки параметров в условиях высокого экспериментального шума снижена на 48% для стрэтчэкспоненциального закона испускания флуоресценции [9].

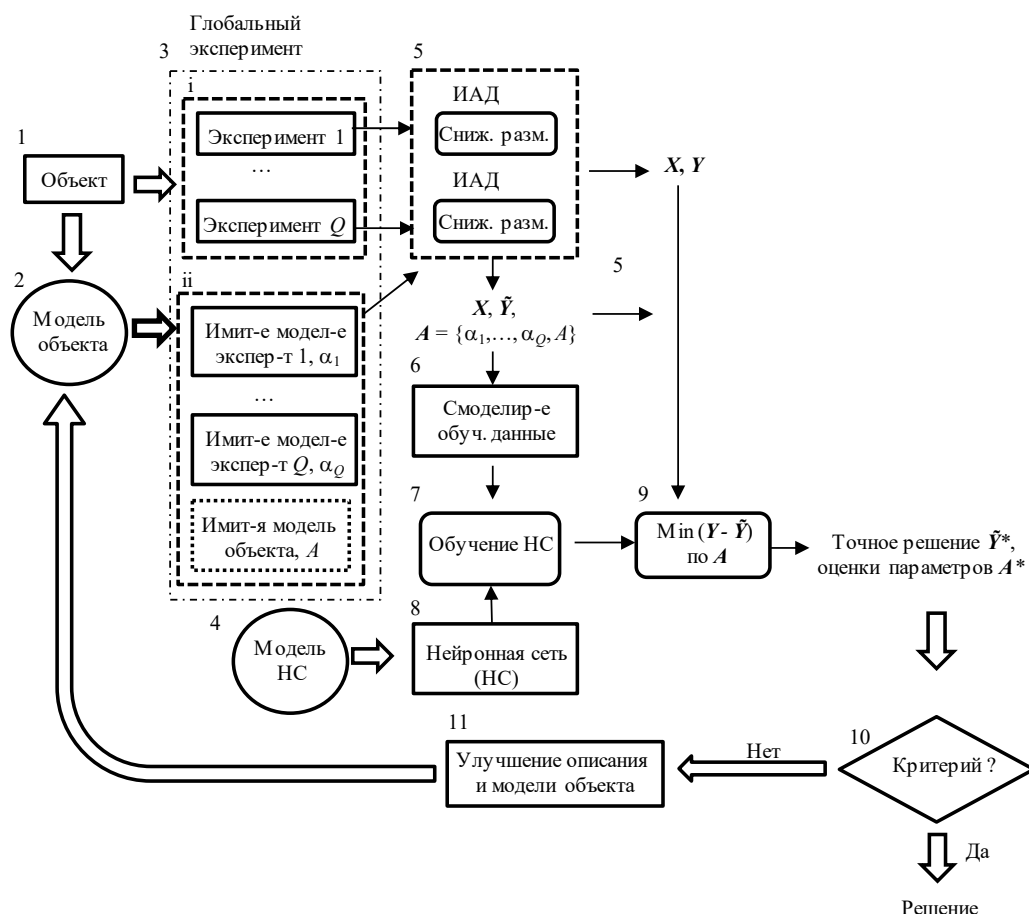


Рис. 1. Схема исследования объекта с использованием Q экспериментов и комплексного подхода

Разработанный комплексный анализ прошел апробацию в ходе исследования флуктуаций интенсивности флуоресценции и построенных на их основе гистограмм числа фотоотсчетов, что позволило определить структурный состав олигомеров белков, повысить точность анализа данных, обеспечить вычислительную производительность, обусловленную высокой скоростью выполнения процедур анализа в сравнении с отдельной обработкой полного набора гистограмм, и представить возможность наглядной визуализации данных в пространстве первых двух главных компонент, что существенно информативнее диаграмм исходных гистограмм [10].

ЗАКЛЮЧЕНИЕ

Разработан комплексный подход на основе методов имитационного моделирования и интеллектуального анализа данных для исследования сложных систем обеспечения прикладной флуоресцентной спектроскопии. Основная фундаментальная задача комплексного подхода – извлечение скрытых знаний в больших биофизических данных, труднодоступ-

ных к обнаружению традиционными методами анализа. Разработанные алгоритмы оценки параметров больших наборов кривых затухания флуоресценции и определения молекулярных агрегатов по данным флуоресцентной флукуационной спектроскопии позволили быстрее и более точно определить параметры биофизических и оптических процессов в смоделированных системах молекулярных соединений, в сравнении с классическим способом анализа, и установить структурный состав олигомеров белков.

БИБЛИОГРАФИЧЕСКИЕ ССЫЛКИ

1. Lakowicz J. R. Principles of Fluorescence Spectroscopy // 3rd ed. New York: Springer, 2006. 954 p.
2. Gryczynski Z. Practical Fluorescence Spectroscopy // Boca Raton: CRC Press, 2020. 792 p.
3. Demchenko A. P. Introduction to Fluorescence Sensing: Volume 1: Materials and Devices // 3rd ed. Cham, Switzerland: Springer, 2020. 673 p.
4. Datta R., Heaster T.M., Sharick J.T. et al. Fluorescence lifetime imaging microscopy : fundamentals and advances in instrumentation, analysis, and applications // J. Biomed. Opt. 2020. Vol. 25(7):071203. DOI: 10.1117/1.JBO.25.7.071203
5. Datta R., Gillette A., Stefely M., Skala M.C. Recent innovations in fluorescence lifetime imaging microscopy for biology and medicine // J. Biomed. Opt. 2021. Vol. 26(7):070603. DOI: 10.1117/1.JBO.26.7.070603
6. Yatskou M. M. Computer Simulation of Energy Relaxation and Transport in Organized Porphyrin Systems // Wageningen, The Netherlands: Ponsen & Looijen Printing Establishment, 2001. 176 p.
7. Яцков Н. Н. Интеллектуальный анализ данных : пособие // Минск: БГУ, 2014. 151 с.
8. Яцков Н. Н., Апанасович В. В. Вычислительная платформа FluorSimStudio для обработки кинетических кривых затухания флуоресценции с использованием алгоритмов имитационного моделирования и интеллектуального анализа данных // Журнал прикладной спектроскопии. 2021. Т. 88, № 3. С. 452-461.
9. Яцков Н. Н., Скакун В.В., Апанасович В. В. Метод обработки кинетических кривых затухания флуоресценции с использованием алгоритмов интеллектуального анализа данных // Журнал прикладной спектроскопии. 2020. Т. 87, № 4. С. 311-322.
10. Яцков Н. Н., Скакун В. В., Недервин-Шипперс Л. И др. Комплексный анализ флукуаций интенсивности флуоресценции молекулярных соединений // Журнал прикладной спектроскопии. 2020. Т. 87, № 4. – С. 628-636.