

Математические методы синтеза моделей биологических систем на основе медицинских данных

М. С. Куприянов¹, А. И. Водяхо², Н. А. Жукова³

СПбГЭТУ «ЛЭТИ»

¹mikhail.kupriyanov@gmail.com, ²aivodyaho@mail.ru,

³nazukova@mail.ru

Д. И. Курапеев¹, М. С. Лушнов²

ФГБУ «НМИЦ им. В. А. Алмазова» Минздрава России

¹dkurapeev@almazovcentre.ru

²lushnov_ms@almazovcentre.ru

В. Ю. Осипов*

Санкт-Петербургский институт информатики и автоматизации РАН

osipov_vasiliy@mail.ru

Аннотация. Отличительной особенностью биологических систем является их высокая сложность. Очень часто эксперты, работающие в области построения моделей биологических систем, не обладают достаточной математической подготовкой, позволяющей строить достаточно сложные корректные модели биологических систем. Авторы предлагают подход к автоматическому синтезу моделей на базе доступных экспериментальных данных. Для решения этой задачи предлагается использовать аппарат относительно конечных многоуровневых операционных автоматов. Описывается алгоритм синтеза таких автоматов. Результаты решения практических задач показывают эффективность предлагаемого подхода.

Ключевые слова: биологические системы; системный анализ; методы синтеза

I. ВВЕДЕНИЕ

Исследование множества свойств биологических систем привело ученых и практических специалистов к необходимости развивать системные подходы для анализа состояния этих систем, оценки происходящих в них изменений. Под системой понимается собранная вместе совокупность элементов, организованных без четкого критерия объединения и упорядочивания [1, 2]. Системный подход в биологии и медицине определяется через свойства и признаки самой системы.

В настоящей работе применён системный подход к моделированию биосистем, подробно изложенный в работе Лушнова А.М. с соавторами. [3]. Применительно к исследованию пациентов приспособительные механизмы можно наблюдать во многих функциональных системах: лейкоцитарной, биохимической, электролитной и других. При анализе изменений в физиологической системе

интерес представляет согласованность внутреннего взаимодействия, т.е. самоорганизации элементов биологической системы.

Основной целью изучения самоорганизующихся систем является оценка их функционального состояния в целом и управление ими с учетом внешних и внутренних влияющих факторов. При этом используются методы системного анализа, математической статистики, многомерного анализа данных. Системный подход к исследованию биологических систем на практике используется очень ограниченно. Поэтому актуальной является задача реализации системного подхода к анализу состояния физиологических систем здоровых людей и пациентов в практике здравоохранения.

II. АНАЛИЗ МОДЕЛЕЙ И МЕТОДОВ СИСТЕМНОГО АНАЛИЗА БИОЛОГИЧЕСКИХ СИСТЕМ

Оценкой функционального состояния организма занимались многие исследователи, предлагавшие различные показатели: физиологические, медицинские, единичные, комплексные, временные и так далее. Большая часть полученных результатов применима для статической оценки отдельных параметров. Очень мало работ по динамическому моделированию этих параметров и практически совсем нет трудов по построению комплексных моделей биологических систем в динамике. Наиболее значимые результаты в этой области получены в ходе многопараметрического моделирования, направленного на исследование функционального состояния организма. Полученные статистические модели, а также методы их построения могут использоваться как основа и руководство для автоматизации процессов моделирования биологических систем.

С позиций решения задач автоматического моделирования, можно выделить следующие особенности биологических систем.

*Работа выполнена по бюджетной теме № 0073-2018-0004 и поддержана РНФ (проект № 17-15-01177).

Первая особенность состоит в том, что анализ и оценка состояния системы осуществляется исходя из изменения значений параметров относительно ранее наблюдавшихся. В настоящее время методиками предусматривается оценка параметров относительно нормативных значений.

Вторая особенность определяется тем, что оценивается согласованность изменений различных параметров системы между собой в контексте общего состояния системы. В случае, если наблюдаемые изменения отражают разладку в функционировании системы, то может анализироваться динамика такого рассогласования.

Для реализации моделей, обладающих приведенными особенностями, возможностей существующих компьютерных моделей биологических систем оказывается не всегда достаточно [4, 5]. Основу традиционных моделей, как правило, составляют регрессионные, качественные и имитационные модели. К основным ограничениям использования этих моделей можно отнести высокую вычислительную сложность, достаточно низкие возможности по их оперативной перестройке.

В [6] выполнен анализ существующих подходов к моделированию сложных объектов, даны их оценки. Предложен новый подход к моделированию, обеспечивающий оперативный синтез моделей объектов по их параметрам. Для формального описания объектов разработаны автоматные модели. Для эффективного применения новых автоматных моделей в интересах исследования биологических систем требуется уточнить эти модели исходя из особенностей предметной области.

III. ПРОБЛЕМА СИНТЕЗА МОДЕЛЕЙ БИОЛОГИЧЕСКИХ СИСТЕМ ПО РЕЗУЛЬТАТАМ ФРАГМЕНТАРНЫХ ИССЛЕДОВАНИЙ ОРГАНИЗМА

Пусть наблюдение за некоторой системой B ведется на интервале времени $\Delta T = [t_s, \dots, t_r]$, где t_s – момент начала наблюдения, а t_r – момент времени, после которого новой информации о системе не поступало. Состояние системы B_s на начальный момент времени не известно. Наблюдение за системой предполагает, что периодически осуществляются управляемые измерения значений отдельных параметров этой системы. В результате становится доступной для анализа последовательность $Z = \{Z_s, \dots, Z_r\}$, где Z_i – результаты наблюдений некоторого множества параметров на i момент времени.

Интерес представляет оценка текущего состояния системы с учетом ранее наблюдавшихся изменений в состоянии системы, а также оценка ожидаемого состояния системы на некотором интервале $[t_{r+1}, \dots, t_{r+k}]$. Оценка должна строиться исходя из имеющихся данных об изменениях $\Delta \hat{B}$, происходящих в системе: $M_r = \{\hat{B}_r | \Delta \hat{B}_{r-1}, \dots, \Delta \hat{B}_{r-j}, Z\}$.

Для комплексной оценки состояния системы требуется определить функционалы, отражающие изменения зависимостей между параметрами. В общем виде

функционалы можно определить как $F = func(\Delta p)$, где $p = \{p_i\}$, $p_i \in P$, P – множество всех измеряемых параметров. При этом интерес представляют устойчивые непротиворечивые зависимости, построенные на множестве функционалов.

Устойчивость во времени определяется продолжительностью интервала времени, на котором зависимость наблюдается $S = \{\Delta t_{s-k} : F(t_s) \& \dots \& F(t_k) = 1\}$. Непротиворечивость определяется исходя из физических свойств параметров. В частности, учитывается, что параметрам биологических систем не свойственны скачкообразные изменения. В связи с этим непротиворечивость оценивается на основе вычисления гладкости кривой: $L = \{\Delta t_{j-k} : curv(\Delta F(t_{i,j})) < \varepsilon\}$.

IV. МОДЕЛИРОВАНИЕ БИОЛОГИЧЕСКИХ СИСТЕМ НА ОСНОВЕ ОТНОСИТЕЛЬНО КОНЕЧНЫХ ОПЕРАЦИОННЫХ АВТОМАТОВ

Аппарат иерархических относительно конечных операционных автоматов [7] позволяет синтезировать иерархические модели объектов на основе исходных результатов измерений параметров.

Структура автоматной модели отражает структуру реальных биологических систем. Нижний уровень содержит информацию о параметрах, более высокие уровни – о подсистемах. Множество допустимых состояний исследуемых систем определяются в терминах изменений относительно предыдущих состояний. Переход между состояниями осуществляется исходя из рассчитываемых значений функционалов. Допустимые множества функций переходов и выходов ограничиваются множеством функционалов, которые могут быть вычислены по имеющимся данным.

Иерархическая автоматная модель определяется как цепочка взаимосвязанных автоматов смежных уровней иерархии

$$DOKA^0 \Leftrightarrow DOKA^1 \Leftrightarrow \dots \Leftrightarrow DOKA^i \Leftrightarrow \dots \Leftrightarrow DOKA^l.$$

Рассмотрим модель i -го уровня на r -ом шаге. В общем виде такая модель описывается как $OKA_r = \{d_{ar}, d_{br}, d_{cr}, F_r^b, F_r^c, DA(d_{br-1}), DB(d_{br-1}), DC(d_{br-1}),$

$FB(d_{br-1}), FC(d_{br-1})\}$, где \bar{d}_{ar} – вектор входных данных;

\bar{d}_{br} – вектор параметров внутреннего состояния; \bar{d}_{cr} –

вектор параметров состояния выхода; F_r^b – функции переходов автомата из одного внутреннего состояния в

другое; F_r^c – функции выходов. Состояния \bar{d}_{br} , \bar{d}_{cr} , \bar{d}_{ar} и функции F_r^b , F_r^c , характеризующие автомат на r -й момент

времени, должны удовлетворять условиям:

$$\bar{d}_{ar} \in DA(\bar{d}_{br-1}), \bar{d}_{br} \in DB(\bar{d}_{br-1}), \bar{d}_{cr} \in DC(\bar{d}_{br-1}),$$

$F_r^b \in FB(\bar{d}_{br-1}), F_r^c \in FC(\bar{d}_{br-1})$. В соответствии с этими условиями, для каждого параметра автоматной модели

определяется множество допустимых значений, которые определяются относительно предыдущего $t-1$ шага.

С учетом особенностей биологических систем автоматная модель представляется в виде $OKA_t = \Delta T_t OKA_{t-1}$, где ΔT_t – оператор, определяющий изменение состояния, произошедшее на интервале $[t-1, t]$. Тогда рассматривая n – кратную итерацию операторов можно системно проследить изменения состояния системы с некоторого момента t_0 : $OKA_t = \Delta T_t \dots \Delta T_1 OKA_{t_0}$.

Поскольку как внутренние состояния, так и выходные данные описывают изменения состояния системы, то нет необходимости разделять множества допустимых значений DB и DC . Можно определить множество $DBC = DB = DC$. Зависимость этих множеств от предыдущего шага сохраняется, т.е. $DBC(b_{t-1})$. Аналогично совпадают множества допустимых функций внутренних переходов и выходов $FBC = FB = FC$. Поскольку множество FBC определяется функционалами, то его можно записать как $FBC(F(b_{t-1}))$.

Предложенная модель биологических систем на основе ОКА является полностью перестраиваемой моделью. Для ее перестройки достаточно переопределить множество вычисляемых функционалов или множество исследуемых состояний биологических систем.

V. РЕКОМЕНДУЕМЫЕ МЕТОДЫ ДЛЯ КОМПЛЕКСНОЙ ОЦЕНКИ БИОЛОГИЧЕСКИХ СИСТЕМ

В области теоретической медицины наблюдается достаточно устойчивая тенденция к сведению многочисленных фактов и законов к нескольким или одному закону. Как правило, при этом используется принцип оптимальности.

Биологический смысл таких моделей заключается в определении максимума корреляционных связей в нормальном состоянии системы и в оценке степени разбалансировки связей при наличии функциональных нарушений. Например, иногда полезно выполнить преобразования, позволяющие выбрать оптимальные подмножества признаков и дать системную оценку для каждого объекта многомерной выборки с использованием критериальной функции, предложенной в [8, 9].

Другой системный статистический метод исследования состояния функциональной системы для выборочной совокупности (массива многомерных данных) основан на определении некоторой оптимизационной суммарной оценки, которая рассматривается как функционал. При этом производится поиск порогов существенности связей и разбиение множества объектов на непересекающиеся классы, которые представляют собой, например, наборы физиологических параметров. [3, 9, 10, 11].

Кроме упомянутых методов могут применяться и другие методы, в частности, оценки функционалов системы на основе матрицы корреляций [11], формирования логических решающих правил [12] и другие.

VI. РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ

Автоматные модели биологических систем синтезировались на основе данных о системе периферической крови (лейкограмма, эритроциты, гемоглобин) и системе кислотно-основного состояния крови – кислотно-щелочного равновесия (КОС-КЩР). Общее число анализируемых параметров составляло несколько десятков параметров. Выборочные данные формировались из 4074 протоколов в основном кардиохирургических вмешательств, анализов КОС (21951 случай), гематологических анализов лейкоцитарной периферической крови (24397 случаев) – лейкограммы (ЛГ) пациентов с кардиопатологией, проходивших лечение в ФГБУ «НМИЦ им. В. А. Алмазова» Минздрава России.

Для вышеперечисленных показателей пациентов функционалы [11] и критериальные функции [9] построены в динамике по упомянутым методикам (рис. 1–рис. 3).

Пример динамической системной оценки (функционала) кислотно-основного состояния КОС и ЛГ в до-, во время и после- операционный периоды на примере проб венозной крови у кардиохирургических пациентов приведен в табл. 1, на рис. 1, 2.

ТАБЛИЦА 1 Динамика функционала КОС венозной крови в до-, во время и после- операционный периоды

Дни	Объём выбор ки	Функ- ционал КОС	Порог сущест- венности связей КОС	Количество классов разбиения параметров КОС
-14	28	13.31	0.42	9
-7	39	26.71	0.27	4
-3	129	12.32	0.36	9
0 - операция	2908	13.98	0.35	8
1	4128	12.95	0.34	9
2	1460	12.22	0.38	9
3	847	10.23	0.47	11
4	535	13.22	0.37	8
5	340	11.71	0.42	9
6	193	11.08	0.46	10
9	215	14.12	0.35	8
14	322	12.17	0.43	9

Наглядное представление описываемых параметров показано на рис. 1, откуда видно, что динамика функционала КОС даёт представление о совокупности метаболических изменений в кислотно-основном обмене пациентов.

Выше приведенные графики системных показателей КОС и ЛГ демонстрируют изменения в периферической крови и в КОС в ходе лечения пациентов, позволяют снизить размерность пространства признаков для облегчения оценки врачом большого количества параметров, компактно описать процесс.

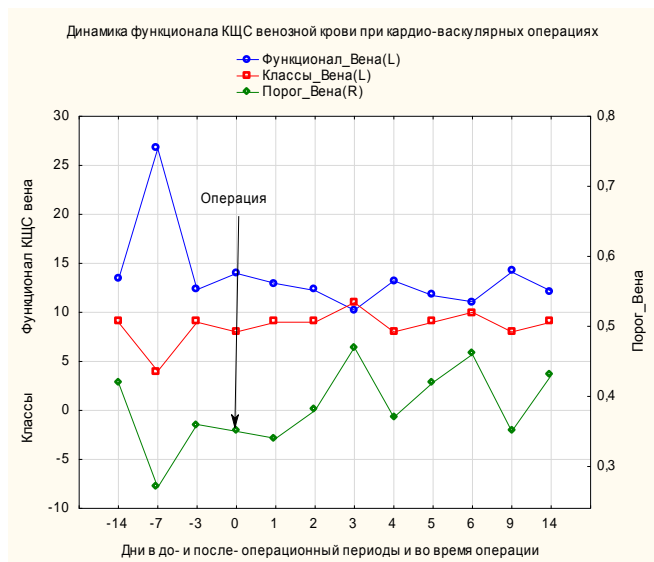


Рис. 1. Динамика функционала КОС, порогов существенности связей и классов разбиения параметров венозной крови у пациентов в до-, во время и после-операционный периоды



Рис. 2. Динамика функционала ЛГ, классов разбиения и порога существенности связей у пациентов в до-, во время и после-операционный периоды

Другой пример – исследовалась система КОС артериальной крови 65 кардиохирургических пациентов, у которых заборы крови проводились 6 раз в ходе лечения. Общая выборка здесь составила 390 наблюдений. Далее для каждого пациента вычислялась критериальная функция (КФ) в динамике. На рис. 3 – динамика КФ КОС пациента Д. Динамика КФ КОС этого пациента, как и других больных указывает на индивидуальный тип реагирования КОС показателей. Поэтому КФ КОС позволят в будущем выявлять типы реагирования индивидуумов на внешние и внутренние воздействия.



Рис. 3. Динамика КФ КОС пациента Д

Полученные результаты экспериментальных исследований позволяют сделать вывод о возможности и применимости предложенных автоматных моделей при моделировании биологических систем.

СПИСОК ЛИТЕРАТУРЫ

- [1] Эшби У.Р. Общая теория систем как новая научная дисциплина // Исследование по общей теории систем. М.: Прогресс, 1969. С. 125-142.
- [2] Анохин П.К. Принципиальные вопросы общей теории функциональных систем // Принципы системной организации функций. М.: Наука, 1973. С. 5-61.
- [3] Лушнов А.М., Лушнов М.С. Медицинские информационные системы: многомерный анализ медицинских и экологических данных. Санкт-Петербург, «Геликон Плюс», 2013. 460 с.
- [4] Haefner J. Modeling Biological Systems: Principles and Applications. Springer US, Springer-Verlag US. 2005.
- [5] Schwartz R. Biological Modeling and Simulation: A Survey of Practical Models, Algorithms, and Numerical Methods (Computational Molecular Biology) 1st Edition. The MIT Press. 2008 Pp. 408.
- [6] Осипов В.Ю. Автоматический синтез программ действий интеллектуальных роботов. Программирование, 2016, № 3. С. 47-54.
- [7] Осипов В.Ю. Синтез результативных программ управления информационно-вычислительными ресурсами // Приборы и системы управления. 1998. №12. С.24-27.
- [8] Osipov V, Zhukova N, Vodyaho A. About One Approach to Multilevel Behavioral Program Synthesis for Television Devices. International journal of computers and communications. Volume 11, 2017, 17-25
- [9] Narendra P.M., Fukunaga K. A branch and bound algorithm for feature subset selection // IEEE Trans. Comput. C. 1977. Vol. 26. P. 917-922.
- [10] Миркин Б.Г. Задачи аппроксимации в пространстве отношений и анализ качественных данных // Автоматика и телемеханика. 1974. № 9. С. 53-61.
- [11] Куперштох В.Л., Миркин Б.Г., Трофимов В.А. Сумма внутренних связей как показатель качества классификации // Автоматика и телемеханика. 1976. № 3. С. 133-141.
- [12] Загоруйко Н.Г., Елкина В.Н., Лбов Г.С. Алгоритм обнаружения эмпирических закономерностей. Новосибирск: Наука, 1985. 112 с.