

РЕЦЕНЗИЯ

върху дисертационен труд за присъждане на образователната и научна степен
“доктор”

Професионално направление: 4.6. „Информатика и компютърни науки”
Научна специалност *“Биоавтоматика”*

Автор на дисертационния труд: инж. Мария Колева Ангелова

Тема на дисертационния труд:
“МОДИФИЦИРАНИ ГЕНЕТИЧНИ АЛГОРИТМИ И ИНТУИЦИОНИСТКИ
РАЗМИТА ЛОГИКА ЗА ПАРАМЕТРИЧНА ИДЕНТИФИКАЦИЯ НА МОДЕЛ
НА ПОЛУПЕРИОДИЧНА КУЛТИВАЦИЯ ”

Рецензент: доц. д-р инж. Иван Симеонов Симеонов

Научно-изследователска група “Математическо моделиране и компютърни
методи” при Института по микробиология на БАН,
Ул. “Акад. Г. Бончев”, бл. 26, София 1113,
Тел.: 979-36-14, E-mail: issim@microbio.bas.bg, issim@abv.bg
GSM: 0889-711-888

1. Актуалност на разработения в дисертационния труд проблем

Разглежданият в дисертацията проблем свързан с параметричната идентификация на нелинеен модел на конкретен биотехнологичен процес (БТП) с използването на т.н. “генетични алгоритми” и на апарата на „интуиционистки рамитата логика” за оценяване на качеството на разработените модификации на генетични алгорими (ГА) е **изключително интересен и актуален**.

Както още във въведението докторантката подчертава, а в Гл. 1 умело и задълбочено доразвива, биотехнологичните процеси включват “жива” система със специфични характеристики, което силно ги отличава от другите обекти за моделиране и управление. Тази “жива” система притежава генетически свойства да се “адаптира, саморегулира, самоуправлява и самопроизвежда”, притежава “памет”, заложена в нейния генетичен механизъм и др.

Изследванията свързани с математическото моделиране, мониторинг, оптимизация и управление на БТП на основата на математически модели водят началото си от 70-те години на 20-ти век. Отначало те са дело главно на химици и биохимици, имащи съответната теоретична подготовка за количествени изследвания и анализи. Учените от областта на автоматиката проявиха по-силен интерес към тази област от средата на 80-те години, като едни от първите в световен мащаб бяха и български учени от бившата ЦЛБА-БАН, някои от които сега са световноизвестни учени. Въпреки големите успехи през 90-те години на 20-ти век по моделиране и управление на БТП, появилите се немалко монографии и стотици статии, в никакъв случай не може да се твърди, че има създадена единна теория в тази сложна интердисциплинарна област.

Заслужава конкретно да се обоснове необходимостта от търсенето и изследването на нови алгоритми за параметрична идентификация на моделите на БТП. Тези модели по същество са нелинейни. С много малки изключения опитите за тяхната линеаризация търпят пълен провал. Но в областта на нелинейните системи въпреки многото изследвания и публикации през последните години теорията все още не е достатъчно развита. Например

едни от сериозните теоретични достижения напоследък са въвеждането на понятията „теоретическа (структурна)” и „практическа” идентифицируемост на моделите на БТП. Първото понятие обуславя съществуването на стриктно решение, т.е. възможност за еднозначно определяне на всички (при глобалната структурна идентифицируемост) или на група от коефициенти (при локалната структурна идентифицируемост) на модела при наличие на „идеални” (в смисъл точни и незашумени) експериментални данни. Въпреки големите трудности при намирането на такива стриктни решения може да се окаже, че „реалните” (зашумени и с определени неточности поради неповтаряемостта на БТП) данни не позволяват да се намери такова стриктно (еднозначно) решение. Т.е. след полагането на значителни усилия от страна на изследователя при определяне на идентифицируемостта на коефициентите на конкретен нелинеен модел на БТП може да се окаже (особено при модели с по-висока размерност), че няма стриктно решение или че има безкрайно много решения. Но е известно, че в такива случаи „инженерният подход” позволява да се намерят решения, които удовлетворяват изследователя без да отговарят на тези стриктни изисквания.

Почти във всички случаи на търсенето на решение на този проблем се решава минимизационна задача. В този смисъл разработването на различни модификации на ГА за целите на параметрична идентификация на модели на БТП е един актуален проблем. Това че за база се използва модел на един конкретен БТП е правилно, защото при нелинейните системи всеки конкретен случай е обект на отделно изследване. При това е избран модел на много интересен БТП, при който един и същи тип микроорганизми (вид дрожди) могат да произвеждат различни продукти в зависимост от метаболитния път на тяхното развитие – при едни условия продуцират хлебна мая, а при други – етилов алкохол.

Заслужава да се отбележи и факта, че в Световната Федерация по Автоматично управление (ИФАК) един от деветте координационни комитети е наречен “Био- екологични системи” и се провеждат регулярни симпозиуми и конференции в тази област.

Въз основа на изложените по-горе съображения може да се направи извода, че **актуалността на дисертационната работа е несъмнена и че в нея са получени оригинални резултати.**

2. Технически данни за дисертационния труд

Дисертационният труд на инж. Мария Колева Ангелова е изложен на 125 страници в увод, 4 глави, заключение (в което са формулирани и приносите и са представени публикациите по дисертацията) и библиография. Резултатите са представени нагледно в голям брой таблици и графики.

3. Оценка на познанията на дисертанта по проблема

В дисертацията е направен задълбочен литературен преглед на известни резултати в няколко научни области.

Литературният обзор включва 129 заглавия, от които 27 са на кирилица и 102 на латиница. Литературните източници на кирилица са 3 книги на руски език (1 книга по оптимално управление на БТП с автор Ю. Станишкис, превърнала се в „класическа” в областта, 1 книга по числени методи и 1 по методи за глобална оптимизация) и 24 източника на български език. 16 от заглавията на кирилица (т.е. около 60 %) са издадени през последните 10 г. Добро впечатление прави цитирането на 3 дисертации и 6 публикации на членовете на научния колектив, в който е работила докторантката.

Литературните източници на латиница са само на английски език. От тях 17 са отпечатани през и след 2010 г., а 29 са отпечатани в периода 2004 – 2009, т.е. 47 от литературните източници на английски език (около 63 %) са отпечатани през последните 10 г.

В литературния обзор са включени и 6 веб страници без достатъчно информация.

Поради факта, че дисертационният труд определено е интердисциплинарен, заслужава да се отбележи и разпределението на използваните литературни източници по отделните научни области:

1. В областта на биотехнологията са цитирани 4 литературни източника на български, изтеглени чрез Интернет. Не бих определил, че това са най-добрите учебници на български в областта, а съществуват и класически книги на английски, макар и не от последните години.

2. В областта на теорията на идентификацията, оптимизацията и числените методи (без генетичните алгоритми) са цитирани 18 заглавия, сред които определено са включени и едни от най-добрите от български автори.

3. В областта на моделирането, идентификацията, оптимизацията и управлението конкретно на БТП (включително и с използване на генетичните алгоритми) са цитирани 49 заглавия, като са включени и едни от най-добрите публикации в тази област.

4. От областта на теорията на генетичните алгоритми, съответните специализирани програмни продукти и няколко приложения на ГА в други области (химия, медицина) са 37 заглавия.

5. От други области (основно интуиционистки размита логика) са цитирани останалите няколко заглавия.

В заключение би могло да се каже, че литературният обзор без да е „претрупан” обхваща по-голяма част от най-новите и значими публикации в основните области на това интердисциплинарно научно изследване. Силно впечатление правят задълбочените интердисциплинарните познания на докторантката, вкл. микробиологията и биохимията на разглеждания биотехнологичен процес, теорията и използването на класическите оптимизационни методи за целите на нелинейната параметрична идентификация, теорията на генетичните алгоритми. Това което ѝ липсва е опит от експериментална дейност поради липсата на такава възможност в звеното, в което е обучавана.

Литературният анализ и съдържанието на представената дисертационна работата показват, че дисертантката познава много добре състоянието на разглеждания проблем, владее съвременната теория на “Биоавтоматиката” и умее да ги използва за решаване на конкретни задачи.

4. Оценка на избраната методика на изследване

При моделирането и управлението на БТП са известни някои обобщени научни резултати, свързани с възможностите за обобщено представяне на техните нелинейни модели, тяхната идентифицируемост, синтез на нелинейни наблюдатели, изследване на устойчивостта и някои методи за управление. Поради изключителната сложност и изменчивост на тези процеси обаче, известните теоретични резултати могат да послужат само като методологична основа за задълбочени изследвания във всеки конкретен случай. Особено труден е проблема за параметрична идентификация на нелинейни модели с голям брой коефициенти при наличие на ограничени (а често и без повтаряемост) експериментални данни.

В този контекст избраната от докторантката методика за използване на „класически” критерий, включващ адитивно четирите променливи на модела, за минимизация на разликата между експерименталните данни и данните от модела при известен интервал на изменение на неизвестните коефициенти с използване на различни модификации на ГА е добра с оглед опита за намиране на „глобално” решение на проблема. Тъй като теоретичното оценяване на валидността на получените решения е изключително трудна задача ми се струва, че само стойността на критерия не е достатъчна. В такива случаи една проста проверка, която би установила валидността на модела (с намерените стойности на коефициентите) в определена област би била верификацията на получените решения с различни експериментални данни

получени например при различни начални условия на променливите X , S и/или pO_2 или на функцията $F(t)$, определяща вида на подхранването.

5. Кратка характеристика на дисертационния труд

Гл. 1 озаглавена „Теоретични основи при параметрична идентификация на модели на ферментационни процеси” в обем от около 30 стр. представлява компилация от знания от няколко научни области. Направен опит да се представят най-важните достижения във всички тези области, което все пак се е отразило негативно основно на частта „Биотехнология”. Но пък описанието на процеса на култивиране на един вид дрожди и неговия математически модел са представени доста пълно и задълбочено, което води до извода, че докторантка има добри теоретични знания в областта. Избрана е известна нелинейна структура на модела на култивиране на дрожди с опит за обобщаването му чрез използване на едно обособено функционално състояние. Моделът включва 4-те основни променливи и понеже е от полупериодичен тип (с подхранване) се описва с 5 ОДУ. За съжаление се разполага само с един набор експериментални данни, при това по мое мнение не особено подходящи за поставената цел. Разгледани са основните понятия в областите параметрична идентификация, методи за оптимизация (вкл. ГА) и интуиционистки размитата логика. След съответните изводи са формулирани стегнато и ясно целите и задачите на дисертацията.

Гл. 2 (в обем от 24 стр.) и Гл. 3 (в обем от 14 стр.) са посветени на параметрична идентификация на избрания нелинеен динамичен модел, като е извършена голяма по обем изчислителна работа и симулационни изследвания при тестването на различни модификации на ГА и сравняване на получените данни от симулационните изследвания с наличния набор от експериментални данни. Направени са сравнения на различните резултати, които са обобщени в съответните изводи.

В Гл. 4 (в обем от 24 стр.) е оценено качеството на представяне на различните алгоритми с използването на апарата на интуиционистки размитата логика. Разработената процедура, особено ако се потвърди твърдението на докторантката, че е универсална и би могла да се приложи и в други случаи, води до заключението за наличие на научни (обобщаващи) приноси на дисертационния труд.

Като Глава 5 е представено заключението, включително формулираните 5 приноса и заглавията на 5-те публикации по дисертационния труд.

6. Основни научни и научно-приложни резултати

В дисертационния труд е представена трудоемка научно-изследователска работа и има приноси, съгласно изискванията за получаване на образователната и научната степен “доктор”. Разработени са голям брой модифицирани генетичните алгоритми за решаване на трудния проблем за параметрична идентификация на коефициенти на нелинейни динамични модели на БТП. Тези приноси според мен са основно с научно-приложен характер и водят до разширяване и подобрене на апарата на ГА.

По-конкретно приносите са:

1. Разработени са единадесет нови модификации на ГА от различен вид (приноси 1 и 2 от формулираните от докторантката), различаващи се по реда на изпълнение на основните генетични оператори, като някои от тях водят до по-добра сходимост на алгоритъма при параметрична идентификация на коефициенти на нелинейни динамични модели на БТП без това да влияе на точността на намереното решение (т.т. на стойността на критерия).

2. Разработена е методика за оптимална настройка на параметрите на ГА от различен вид и процедура за целенасочен генезис на параметрите на модела (приноси 3 и 4 от формулираните от докторантката). Приложението им при параметрична идентификация на математически модел на полупериодичен БТП (култивиране на дрожди) води до значително намаляване (до 40 %) на времето за сходимост на алгоритмите.

3. Разработена е процедура с приложение на интуиционистки размита логика (нов теоретичен апарат с все още малко на брой конкретни приложения) за оценка качеството на представяне на ГА. Процедурата е успешно приложена при параметрична идентификация на математически модел на полупериодичен БТП (култивиране на дрожди) за оценка на различни видове ГА. Предложената процедура е универсална и може да бъде успешно приложена както при параметрична идентификация на модели на различни обекти, така и за други алгоритми за оптимизация.

Последният принос, според мен, е с елементи и на научен такъв.

7. Оценка на личния принос на дисертанта при разработване на дисертационния труд

От прочитането на дисертационния труд и направените публикации по него и от разговорите с инж Мария Колева Ангелова оставам с твърдото убеждение, че това е труд изпълнен лично от нея при методичната помощ на научните си ръководители.

8. Оценка на публикациите по дисертацията

По дисертационния труд са представени **5 отпечатани публикации**. Една от тях е в Доклади на БАН, като за годината на отпечатване (2012) това наше реномирано списание е имало $ИФ = 0.211$. Друга статия на докторантката е отпечатана в едно сравнително ново международно научно списание (Int. J. of Chemical Engineering). Един доклад изнесен на международен научен форум (Eighth IMACS Seminar on Monte Carlo Methods, August 29 – September 2, 2011, Borovets, Bulgaria) е отпечатан като глава от книга издадена от немско издателство, а останалите две представени публикации по дисертацията представляват доклади, изнесени на български научни форуми. И петте представени публикации са в съавторство с единия от ръководителите на дисертацията, като докторантката е първи автор във всяка от тях.

Известни са ми 9 цитирания на четири от представените публикации по дисертацията, както и 36 цитирания на общо 9 научни публикации на докторантката.

Наукометричните показатели на докторантката отговарят на изискванията (а по някои от показателите ги надхвърлят) на Правилника за придобиване на научни степени в ИБФБМИ – БАН.

Всичките публикации са по темата и отразяват напълно приносите в дисертационния труд.

Въз основа на горното **считам, че получените резултати са получили широка публичност както у нас, така и в чужбина.**

9. Оценка на Автореферата

Представеният автореферат (в обем от 44 стр.) е изготвен в съответствие със съществуващите изисквания и отразява напълно основните резултати и постижения в дисертационния труд. Известни трудности при сравняване с дисертационния труд представлява несъответствието в номерацията на фигури и таблици. Като недостатък отчитам и липсата на резюме на английски език.

10. Забележки, препоръки и въпроси към дисертационния труд

Забележки и препоръки:

1. Един от основните етапи на моделирането (както е указано и от докторантката) е верификацията (валидацията) на модела. Тъй като параметричната идентификация на нелинейния модел на процеса е етап на моделирането, логически следва, че най-важният критерий за адекватността на модела ще бъде неговата валидност за други експериментални данни, получени при стойности на някои от основните параметри различни от тези, за които е направена параметричната идентификация. В случая такава проверка би било най-лесно да

се направи с данни получени при други начални стойности на концентрацията на биомасата (инокулом), на субстрата (глюкозата), и/или на разтворения кислород. Би могло да се направи и при друг вид на функцията на подхранването или на концентрацията на субстрата в подхранващия разтвор. При по-голям брой експерименти би могло да се направи заключение и за областта на валидност на модела със съответния(ите) набор(и) от стойности на коефициентите в него. За това обаче е необходимо да се разполага със съответната експериментална база и квалифициран персонал.

2. Масообменният коефициент $\kappa_L^{O_2} a$ е характеристика на съответния биореактор и би могъл лесно да се определи предварително, с което ще се намали с 1 броят на оценяваните коефициенти. В общ случай $\kappa_L^{O_2} a$ е променлива величина, но при промяна на оборотите на разбъркване и/или дебита на подавания въздух (което не е изпълнено в конкретния случай, защото те са поддържани постоянни).

3. От всички фигури с експериментални данни се вижда, че концентрацията на субстрата (глюкоза) е права линия, т.е. очевидно има автоматично регулиране в затворен контур по тази променлива. Въпреки че при симулиране на модела след определяне на съответния набор от коефициенти моделът се държио нелошо, считам че такива експериментални данни не са подходящи за целите на параметричната идентификация. Моята препоръка е да се работи с данни от чисто периодичен процес, като след това данните от процес с подхранване да служат за верификация.

4. Не считам за убедително твърдението, на основата на посочените резултати от постъпковата идентификация с "класически" методи на стр. 43 и 44, че традиционните методи за оптимизация не се справят адекватно с проблема по следните две причини:

А) Ако бе избран оптимизационен метод с ограничение нямаше да се получи отрицателна стойност на един от коефициентите. Друг е проблемът, че е много вероятно решението да спре на ограничението.

Б) Наборът от параметри p_1 е неудачен, защото се съдържат едновременно коефициентите μ_{2S} и k_S (единият от тях е в числителя, а другият е в знаменателя на нелинейната функция за специфичната скорост), а самата докторантка показва (стр. 26), че в такъв случай не могат да се получат добри резултати.

5. Има някои очевидни стилни и печатни грешки – напр. на стр. 107 се касае за Табл. 4.20 вместо 4.19, на стр. 69 очевидно се касае за точност на апроксимацията, а не на модела.

Считам, че повечето от тези забележки нямат особено съществен характер, не намаляват положителните качества на дисертационния труд и ще послужат за бъдещата научна работа на докторантката.

Въпроси:

1. Как се измерва концентрацията на етанола в културалната среда?

2. Какъв е вида на функцията на подхранване $F(t)$?

3. Възможно ли е инхибиране на този процес по глюкоза и/или по етанол и как би се отразило това на модела?

4. Възможно ли е управление на разтворения кислород (pO_2) в биореактора при експериментите и как може да става това?

5. При всички експериментални данни се вижда, че в 12-тия и в 13-ия час от началото на експеримента концентрацията на етанола се повишава, което не може да се апроксимира добре с нито един набор от коефициенти. Това може да се дължи на следните две причини:

А) данните са погрешни;

Б) започнал е синтез на етанол и оттам неговата концентрация се повишава.

Кое е вярното?

6. Какво е n_i в критерия (2.1) и какви са неговите стойности? Правени ли са опити да се отчетат различните размерности на променливите в този критерий?

7. Възможно ли е при ГА да се зададе интервал на началните стойности на идентифицираните параметри?

12. Лични впечатления от докторантката

Познавам докторантката от съвместни участия на научни форуми и от участието ми в изпитна комисия за провеждане на кандидатския ѝ минимум. От запознаването ми с документите съм информиран, че тя е завършила магистърска степен в ТУ-София, специалност „Биоелектроинженерство” с успех от курса на обучение мн. добър и е защитила с отличие дипломната си работа през 2006 г. Взела е всички полагащи се изпити по време на редовната си докторантура, като по кредитната система за оценяване на докторантите на БАН е събрала 526 т. при задължителен минимум от 250 т.

Макар че през последните години не следя отблизо нейното научно развитие, запознаването ми с дисертацията и публикациите ѝ потвърждават мнението ми, че тя е задълбочен и сериозен млад учен с големи перспективи за развитие в областта на теорията и практиката на научната специалност “Биоавтоматиката”.

13. Заключение

Представената дисертационна работа показва, че инж. Мария Колева Ангелова познава и владее професионално съвременните основи на моделирането и нелинейната идентификация на БТП (включително на съвременния софтуер за тези цели, микробиологията и биохимията на моделираните процеси) и особено апарата на т.н. “генетични алгоритми”, може да прилага успешно теоретичните си познания при решаването на конкретни задачи и да извършва самостоятелни научни изследвания в указаната област.

Темата на работата е интересна и актуална. Работата е добре структурирана, притежава сериозна литературна справка, извършена е трудоемка научно-изследователска работа и са получени оригинални научно-приложни приноси, които са отразени в необходимите публикации. Всичко това представя дисертантката като оформен млад учен в областта на научната специалност “Биоавтоматиката”.

Въз основа на горното считам, че дисертационната работа отговаря на изискванията на Закона за развитието на академичния състав в република България, на Правилника за прилагане на закона за развитието на академичния състав в република България и на Правилника за придобиване на научни степени в ИБФБМИ – БАН и предлагам убедено да се присъди образователната и научна степен степен “Доктор” на инж. Мария Колева Ангелова.

20.06.2014 г.

София

Рецензент:

(доц. д-р инж. Иван Симеонов)