

БЪЛГАРСКА АКАДЕМИЯ НА НАУКИТЕ
ИНСТИТУТ ПО БИОФИЗИКА И БИОМЕДИЦИНСКО ИНЖЕНЕРСТВО

**МОДИФИЦИРАНИ ГЕНЕТИЧНИ АЛГОРИТМИ И
ИНТУИЦИОНИСТКИ РАЗМИТА ЛОГИКА ЗА ПАРАМЕТРИЧНА
ИДЕНТИФИКАЦИЯ НА МОДЕЛ НА ПОЛУПЕРИОДИЧНА КУЛТИВАЦИЯ**

инж. Мария Колева Ангелова

АВТОРЕФЕРАТ

на дисертационен труд
за присъждане на образователна и научна степен „доктор”
професионално направление 4.6. „Информатика и компютърни науки“
(научна специалност „Биоавтоматика“)

Научни ръководители: проф. д-н Стоян Цонков
доц. д-р Таня Пенчева

СОФИЯ
2014

Дисертационният труд е обсъден и допуснат до защита на разширено заседание на секция „Биоинформаика и математическо моделиране“ на ИБФБМИ-БАН, състояло се на 31.03.2014 г.

Дисертационният труд съдържа 125 страници, в които 19 фигури, 50 таблици и библиография, включваща 129 заглавия.

Защитата на дисертационния труд ще се състои на2014 г. от ч. в залата на ИБФБМИ-БАН, ул. „Акад. Георги Бончев“, бл. 105 на открито заседание на научното жури в състав:

1. проф. дтн Стоян Стоянов
2. доц. д-р Иван Симеонов
3. доц. д-р Сотир Сотиров
4. доц. д-р Стефка Фиданова
5. доц. д-р Таня Пенчева

Материалите по защитата са на разположение на интересувашите се в канцеларията на ИБФБМИ-БАН, ул. „Акад. Георги Бончев“, бл. 105.

Автор: инж. Мария Колева Ангелова

Заглавие:

Модифицирани генетични алгоритми и интуиционистки размита логика за параметрична идентификация на модел на полупериодична култивация

ОБЩА ХАРАКТЕРИСТИКА НА ДИСЕРТАЦИОННИЯ ТРУД

Актуалност: Едно от най-значимите постижения на научно техническия прогрес през последните десетилетия е разработването на научна основа на нови производствени процеси, в които се използват принципите на биологичните реакции. Биотехнологията, като един от най-високо ефективните отрасли на производството се оценява с основание като технология на бъдещето. Получаването на важни биологично-активни вещества, посредством използване на продуктите от жизнената дейност на микроорганизми, вируси, растителни или животински клетки и техни компоненти, намира приложение в различни области на икономиката, бита и научните изследвания и оформя едно от най-перспективните направления в развитието на биотехнологията. Тази сложна интердисциплинарна научно и научно-приложна област, изгражда и развива своите методи на основата на най-съвременните постижения в областта на молекулярната биология, генетиката, генното инженерство, биохимията, микробиологията, физиокохимията, електрохимията, инженерната химия и др.

Поради широкото приложение на биотехнологичните процеси (БТП) в различни отрасли на индустрията, проблематиката, свързана с тяхното изследване е изключително актуална. Изучаването на БТП и по-специално на ферментационните процеси (ФП) като основна част от БТП е свързано с редица предизвикателства не само, защото свойствата на ФП до голяма степен се определят от свойствата на техните биологични компоненти, но и поради факта, че тези процеси се различават качествено от протичащите в неживата природа. ФП са сложни, динамични системи с взаимозависими и променящи се във времето процесни променливи, което от своя страна води до описанието им с нелинейни модели с много сложна структура. Ето защо тяхното моделиране и висококачествено управление не е нито лесна нито тривиална задача. За разработването на един адекватен математически модел е необходимо първо да бъде избран подходящ подход за моделиране, след което да бъдат използвани ефективни методи за параметрична идентификация на модела. Незадоволителните резултати, до които достигат традиционните методи за оптимизация при параметрична идентификация на модели на ФП провокират идеята, като алтернатива да бъдат приложени еволюционни алгоритми. Измежду тях генетичните алгоритми (ГА), вдъхновени от теорията на Дарвин за еволюция на видовете, са се доказали като изключително подходящи за параметрична идентификация на различни по своята същност и специфика задачи, и особено за параметрична идентификация на ферментационни процеси.

Предметът на дисертационния труд е свързан с разработване и прилагане на ефективни модификации на едно- и многопопулационни генетични алгоритми както и на надеждни процедури за оценка качеството на представяне на тези алгоритми, при параметрична идентификация на модел на реален полупериодичен процес за култивация на дрожди *S. cerevisiae*.

Публикации и апробация на резултатите в дисертационния труд.

Основните резултати в дисертационния труд са отразени в 5 публикации: 2 в специализирани списания, 1 на международна конференция, проведена в България и 2 на национални конференции.

Технически данни. Дисертационният труд е в обем от 125 страници – увод, четири глави, заключение, авторски публикации по темата и списък на използваната литература, съдържащ 129 литературни източника. В автореферата е използвано собствено номериране на формулите, фигурите и таблиците.

КРАТКО СЪДЪРЖАНИЕ НА ДИСЕРТАЦИОННИЯ ТРУД

Глава 1

Теоретични основи при параметрична идентификация на модели на ферментационни процеси

Биотехнологиите заемат водещо място в съвременното научно и технологично развитие. Като технологии, които използват живи микроорганизми, биологични системи или техни производни за целево създаване и модифициране на продукти и процеси, те намират широко приложение в редица индустриални области - от земеделието до медицината и тежката промишленост. Поради това биотехнологията и свързаните с нея технологии имат голямо социално-икономическо значение [4, 7-9].

Едно от основните направления в съвременната биотехнология е микробиологичният синтез [4, 7-9]. Той се свързва не само с получаването на целеви продукти от жизнената дейност на микроорганизмите, но и с изучаване на закономерностите, присъщи на популациите от микроорганизми, култивирани в изкуствено създадени условия. Производствените процеси, които водят до получаването на продукти от жизнената дейност на микроорганизмите са известни като биотехнологични процеси. Основен етап на голяма част от тези процеси е ферментационният стадий (т. нар. ферментационен процес), свързан с култивиране, биосинтез, биотрансформация, биоконверсия и др.

Установяването на закономерностите в развитието на микробналната култура служи като научна основа за създаването на специфични системи за контрол и управление на ФП. Реализирането на управляем микробиологичен синтез не е лесноизпълнима задача, тъй като трябва да се отчете факта, че се управлява биологична (жива) система и като такава тя се характеризира с редица специфични особености. Висококачественото управление на ФП може да бъде обусловено от създаването на адекватни математически модели, които отчитат спецификата на тези процеси.

Особености при моделирането на ферментационни процеси

При реалните процеси, каквито са ФП, се отличават два основни подхода на моделиране, разгледани подробно в [1, 6, 8, 9]. Единият се основава на известни физикохимични и/или биохимични закономерности (аналитичен). Вторият подход (експериментален) се основава само върху експериментални данни, които се обработват чрез различни математически методи, без да се разглежда качествената страна на изучавания обект.

На сегашното ниво на познания за микробиологичния синтез най-резултатен е комбинираният подход за построяване на математически модел, т. е. конкретния вид на зависимостите се намира аналитично, а числените значения на коефициентите се търсят по експериментални данни.

Моделите на ФП се изграждат при следните основни допускания:

- В биореактора са осигурени „идеални“ условия на смесване;
- Скоростите на отделните микро и макро реакции са такива, че мигновено се реагира на измененията на въздействащите върху процеса фактори.

Моделиране на процес на полупериодична култивация на дрожди

Дрождите са важни микроорганизми с разнообразно приложение в различни области на човешката дейност. Активността на дрождевите клетки се използва за промишлени цели при производството на хляб, бира и вино. Значително приложение дрождите намират и в модерните биотехнологии, в сферата на здравеопазването, в генното инженерство, при производството на протеин и опазването на околната среда.

Описание на процес на полупериодична култивация на дрожди *S. cerevisiae*

В настоящото изследване са използвани експериментални данни от полупериодична култивация на дрожди *S. cerevisiae* H620, проведена в Института по техническа химия в ХанOVER, Германия. Култивацията на дрождите е осъществена в 1.5 l реактор, използвайки среда на Шацман [25]. Концентрацията на глюкозата в подхранващия разтвор е 50 g/l. Температурата се поддържа на 30°C, pH – на 5.7. Скоростта на разбъркване е 500 rpm.

Глюкозата е основен субстрат на така описаната полупериодична култивация на дрожди, а като продукти от ферментацията се получават етанол и въглероден диоксид. Освен продукт, при определени условия, когато в средата няма наличие на глюкоза и концентрацията на разтворен кислород е над критичното ниво, етанолът може да бъде и субстрат.

Математически модел на полупериодична култивация на дрожди *S. cerevisiae*

Математическият модел, получен на базата на баланс на масите и описващ скоростите на клетъчен растеж, консумацията на субстрат (глюкоза), формирането на етанол и потреблението на кислород за гореописания процес на полупериодична култивация на дрожди *S. cerevisiae*, се представя в следния вид [25, 34]:

$$\frac{dX}{dt} = \mu X - \frac{F}{V} X \quad (1)$$

$$\frac{dS}{dt} = -q_S X + \frac{F}{V} (S_{in} - S) \quad (2)$$

$$\frac{dE}{dt} = q_E X - \frac{F}{V} E \quad (3)$$

$$\frac{dO_2}{dt} = -q_{O_2} X + k_L^{O_2} a (O_2^* - O_2) \quad (4)$$

$$\frac{dV}{dt} = F \quad (5)$$

където S , X , E и O_2 са непрекъснати и диференцируеми функции, описващи съответно концентрациите на субстрат [g/l], биомаса [g/l], етанол [g/l] и

разтворен кислород [%]; O_2^* – концентрация на насищане с разтворен кислород [%]; $k_L^{O_2} a$ – масообменен коефициент [1/h]; S_{in} – концентрация на субстрат в подхранващия разтвор [g/l]; μ , q_S , q_E , q_{O_2} – специфични скорости на растеж/консумация съответно на биомаса, субстрат, етанол и разтворен кислород [1/h]; F – скорост на подхранване [l/h]; V – обем на биореактора [l].

Zhang и колектив [34] предлагат подход за моделиране на култивационни процеси на дрожди *S. cerevisiae*, съгласно който текущите състояния на системата се дефинират като функционални състояния. На базата на този подход, за разглежданата тук култивация на дрожди са налице условия за разпознаване на едно функционално състояние (ФС), известно като *второ функционално състояние* (или състояние на смесено окисидиране) [34]. Процесът навлиза в това състояние, когато концентрацията на захар намалее до или спадне под критичното ниво, но все още има достатъчно разтворен кислород в средата. Специфичните скорости на растеж/консумация за състоянието на смесено окисидиране, се описват със следните локални модели [34]:

$$\mu = \mu_{2S} \frac{S}{S + k_S} + \mu_{2E} \frac{E}{E + k_E},$$

$$q_S = \frac{\mu_{2S}}{Y_{SX}} \frac{S}{S + k_S}, \quad q_E = -\frac{\mu_{2E}}{Y_{EX}} \frac{E}{E + k_E}, \quad q_{O_2} = q_E Y_{OE} + q_S Y_{OS}, \quad (6)$$

където μ_{2S} , μ_{2E} са максималните стойности на скоростите на растеж, съответно по субстрат и по етанол [1/h]; k_S , k_E – константи на насищане по субстрат и етанол [g/l]; Y_{SX} , Y_{EX} , Y_{OE} , Y_{OS} – икономични коефициенти [g/g].

В този случай специфичната скорост на растеж се представя като сума от две събираеми, описващи дрождевия растеж съответно за сметка на консумацията на субстрат и за сметка на консумацията на етанол. И двете събираеми имат структурата на кинетика на Моно, която се използва и за специфичните скорости на консумацията на етанол и на субстрат. Специфичната скорост на консумация на кислород се представя като сума от пропорционални отношения спрямо специфичните скорости на консумацията на етанол и на субстрат.

Всички променливи на процеса се описват с непрекъснати и диференцируеми функции като параметрите на моделите приемат строго положителни стойности.

Параметрична идентификация на модел на ферментационен процес

Характерните особености на ФП правят математическото им моделиране изключително трудно [1, 5, 6, 8, 9]. Ето защо се изисква особено внимание и специфичен подход в процеса на моделиране на ФП, както и използване на ефективни методи за последваща параметрична идентификация на разработените модели. Оценката на параметрите на даден модел може да се разглежда като оптимизационен проблем, тъй като по същество представлява търсене на „оптимални“ стойности на параметрите на базата на снети експериментални данни. Методите за оценка на параметрите в моделите биват традиционни и стохастични.

Нерешени проблеми при традиционните методи за идентификация на модели на ФП си остават силната зависимост на решенията от началните условия на процедурите и попадането им в локален екстремум, а при методите без ограничения и намирането на некоректни, от биологична гледна точка, стойности на параметрите на модела. Времето за изчисление при някои класове детерминистични методи се увеличава значително (често експоненциално) с увеличаване на размерността на оптимизационната задача, други методи са чувствителни към параметрите на самия алгоритъм и главно към вида и сложността на целевата функция.

Стохастичните методи, от своя страна, преодоляват проблема с по-голямата размерност на решаваните задачи и дават задоволителни резултати за относително кратко време [3].

Същност на генетичните алгоритми

Генетичните алгоритми като стохастичен метод за оптимизация са директна, случайна търсеща техника за намиране на глобално решение в комплексното многомерно пространство на търсене [3, 12, 15, 21]. От биологична гледна точка ГА са базирани на механиката на естествената селекция и генетика, съобразно еволюционната теория на Дарвин. В термините на генетичните алгоритми [15] хромозомите представляват набор от гени, които кодират неизвестните променливи. Всяка хромозома представлява допустимо решение на поставената задача. Множеството от различни хромозоми, съставя текущо поколение, върху което се прилагат еволюционни оператори като селекция, кръстосване и мутация за достигане до ново поколение [15, 27].

Еднопопулационен генетичен алгоритъм

Стандартният еднопопулационен генетичен алгоритъм (СГА), търси глобален оптимум на решавания проблем като работи с една популация от индивиди [15]. При стартиране СГА създава популация от n кодирани хромозоми (индивиди), генерирани на случаен принцип. Всяка от тези изкуствени хромозоми се състои от низове (или гени) с определена дължина (брой елементи) и всеки ген съдържа информация за съответния параметър. След това алгоритъмът изчислява стойността на целевата функция и функцията на изява („фитнес“ стойности) на всеки индивид в началната популация. Съобразно получените „фитнес“ стойности СГА извършва подреждане и селектиране на най-подходящите индивиди като хромозомите с по-голяма функция на изява имат по-голям шанс да бъдат избрани и възпроизведени. Селектираните индивиди първо се подлагат на *кръстосване*, със зададена вероятност. При *кръстосването* гените от родителите се комбинират с цел да се формира изцяло нова хромозома. След това, с цел да се предотврати попадането на решението в локален оптимум и за да се разнообрази популацията се извършва операторът *мутация*. В термините на ГА *мутация* означава произволна промяна на стойността на ген в хромозомата. По този начин се получават нови индивиди, като стремежът е те да наследят възможно най-добрата комбинация от характеристики на своите родители. За новополучените индивиди се изчислява целевата функция и функцията на изява и се извършва замяна на част от старите индивиди с нови. Така получената популация се използва в следващия цикъл на СГА.

Еволюцията на популацията продължава докато се удовлетвори някой от критериите за прекратяване на алгоритъма, като това може да бъде изпълнението на определен брой генерации, достигането на предварително зададена точност на решението и др.

Многопопулационен генетичен алгоритъм

Стандартният многопопулационен генетичен алгоритъм (МПГА), в сравнение с еднопопулационния генетичен алгоритъм, е по-близък до процесите, протичащи в живата природа, тъй като при него независимо една от друга се развиват няколко популации, наречени субпопулации.

След определен брой генерации (време на изолация), част от индивидите мигрират между субпопулациите.

Основни понятия от интуиционистки размитата логика.

Законът за изключеното трето гласи, че едно съждение е или истина, или е лъжа. Трета възможност няма. Това е така, доколкото се разглежда класическата логика. Размитата логика (*fuzzy logic*) е форма на многовариационна логика, произлизаща от тероията на размитите множества с цел да отрази това, което е относително, а не точно. В сравнение с множествата от двоичната логика, размитата логика борави с променливи, които могат да бъдат със стойности между 0 и 1 и не е ограничена единствено до истинни стойности (истина = 1 и лъжа = 0), така както в класическата логика [18]. По-модерна (поради близостта ѝ с реалните ситуации) е интуиционистки размитата логика (ИРЛ), която дава още една степен на свобода чрез въвеждането на допълнителен процент на несигурност. Така сборът от процентите лъжа/истина, с които се оценява вярностната стойност на съждение, често може да се окаже и строго по-малка от единица. Интуиционистки размитата логика се изгражда на базата на интуиционистки размитите множества, предложени от К. Атанасов през 1983 г. [10, 11], като разширение на теорията на размитите множества на Л. Заде.

Апаратът на интуиционистки размитата логика намира все по-широко приложение в различни области на науката, като едновременно с това се търсят възможности за неговото ефективно приложение в редица други научни направления.

ИЗВОДИ

Въз основа на представения анализ могат да бъдат обобщени следните по-важни изводи:

1. Дрождите са важни микроорганизми с широко приложение в различни отрасли на промишлеността. Това обуславя техния избор като обект за изследване в настоящата дисертация.
2. Ферментационните процеси са живи системи и се различават качествено от процесите протичащи в неживата природа. Поради ясно изразената им неопределеност, нестационарност и неповтаряемост тези процеси се описват с нелинейни модели, с много сложна структура.
3. Параметричната идентификация на модели на ФП е процедура, която поставя редица предизвикателства. Най-важното условие

за разработването на адекватен модел е свързано с надеждно оценяване на неговите параметри.

4. Генетичните алгоритми, поради многобройните си предимства са се доказали като подходяща алтернатива на класическите оптимизационни методи за параметрична идентификация на модели на ФП.
5. Стремехът към усъвършенстване на генетичните алгоритми е основна предпоставка за разработване на различни техни модификации и провокира търсенето на нови такива. От тук произтича необходимостта от създаване на допълнителни надеждни методи за оценяване качеството на представяне на разработените ГА.

Цели и задачи на дисертационния труд

Въз основа на извършения анализ и направените изводи основните цели, които се поставят в дисертационния труд са:

1. Изследване на генетични алгоритми и техни модификации за параметрична идентификация на модел на култивация на дрожди *S. cerevisiae*.
2. Изследване приложимостта на апарата на интуиционистки размитата логика за оценка качеството на представяне на разработените ГА.

За постигане на поставените цели са дефинирани следните задачи:

1. За параметричната идентификация на модел на полупериодична култивация на дрожди да се изследват:
 - 1.1. Традиционни оптимизационни методи
 - 1.2. Генетични алгоритми, включително:
 - 1.2.1. Разработване на модификации на еднопопулационен генетичен алгоритъм;
 - 1.2.2. Разработване на модификации на многопопулационен генетичен алгоритъм;
 - 1.2.3. Оптимална настройка на параметрите на генетичните алгоритми.
2. Изследване възможността за повишаване точността на разглеждания модел за полупериодична култивация на дрожди чрез:
 - 2.1. Разработване на процедура за целенасочен параметричен генезис.
 - 2.2. Приложение на разработената процедура при параметрична идентификация на модел на полупериодична култивация на дрожди *S. cerevisiae* с еднопопулационни и многопопулационен ГА.
3. Сравняване качеството на представяне на ГА, чрез използване на интуиционистки размита логика.
 - 3.1. Разработване на процедура за сравняване качеството на представяне на различни ГА с прилагане на апарата на ИРЛ.

3.2. Приложение на процедурата за сравняване качеството на представяне в задачата за параметрична идентификация на модел на полупериодична култивация на *S. cerevisiae* при:

3.2.1. различни видове еднопопулационни генетични алгоритми;

3.2.2. различни видове многопопулационни генетични алгоритми;

3.2.3. стандартен СГА и стандартен МПГА;

3.2.4. ГА с различни стойности на GGAP.

Глава 2

Параметрична идентификация на модел на полупериодична култивация на дрожди *S. cerevisiae*

В практиката не съществува универсален метод за решаване на всички оптимизационни задачи, затова за целите на параметричната идентификация на модел на полупериодична култивация на дрожди *S. cerevisiae* са тествани възможностите на някои от най-широко използваните класически оптимизационни методи, разработени в *Optimization Toolbox* на *Matlab* 7.

В *Matlab* средата е разработен *Simulink* модел за изследвания процес, който описва конкретните диференциални уравнения и специфичните скорости на растеж или консумация. За да се определят кинетичните параметри за *Simulink* модела и за всеки оптимизационен метод, е разработен код, съдържащ необходимите инструкции за начални стойности на параметрите, граници на изменението им (където е допустимо), както и за параметрите на самия оптимизационен метод.

Като критерий в разглежданата тук многокритериална оптимизационна задача, за всички извършени оптимизационни процедури е използвана стойността на средно-квадратичното отклонение между изхода на модела и експерименталните данни, получени по време на култивацията. Стойността на критерия е нормирана, като е отчетена и различната размерност на променливите на вектора на състоянието. Критерият се представя във вида:

$$J = \sum_{i=1}^4 \sum_{j=1}^{n_i} \left(\frac{Y_{ij}^* - Y_{ij}}{Y_{ij} \cdot n_i} \right)^2 \rightarrow \min, \quad (7)$$

където Y_{ij}^* са моделно предсказаните данни, $Y_{ij} = [X, S, E, O_2]$ са експерименталните данни, а n_i е размерността на променливата на състоянието.

Параметрична идентификация с класически методи за оптимизация

Оценяването на деветте параметъра на изследвания модел (1) – (5) и специфични скорости (6) на полупериодичната култивация на дрожди *S. cerevisiae* е извършено последователно с всеки един от четирите изброени по-долу традиционни метода за оптимизация:

- Метод на деформируемия симплекс (*Nelder-Mead Simplex method*) – използва се функцията “*fminsearch*”;

- **Минимаксен метод (Minimax method)** – използва се функцията “fminimax”;
- **Квазинютонови методи:**
 - *Levenberg-Marquardt* – използва се функцията “lsqnonlin” и
 - *Gauss-Newton* – използва се функцията “lsqnonlin” с допълнителна настройка в опциите *LevenbergMarquardt*, 'off' .

Границите на изменение на оценяваните параметри са определени въз основа на литературни източници [17, 19, 20, 23-26, 31-33] и са представени съответно в Таблица 1. Обобщаването на достоверни интервали на изменение на параметрите, обаче е възпрепятствано от спецификата на разглеждания в дисертацията ферментационен процес и липсата на достатъчно данни от параметрична идентификация. Пълен набор от стойности за всички оценявани девет параметъра е наличен единствено в [25]. Специално за параметър Y_{OS} , поради липсата на информация, границите се определят като $\pm 50\%$ от представената в [25] стойност.

Таблица 1 Граници на изменение на параметрите

Параметър	Граници на изменение	
	Долна граница	Горна граница
μ_{2S} , 1/h	0.17	0.95
μ_{2E} , 1/h	0.01	0.34
k_S , g/l	0.1	0.5
k_E , g/l	0.1	5.01
Y_{SX} , g/g	0.05	0.65
Y_{EX} , g/g	0.16	2.08
$k_L^{O_2} a$, 1/h	55	1200
Y_{OS} , g/g	399	1199
Y_{OE} , g/g	0.0013	8340

Поради голямото влияние на началните условия върху получените решения, класическите оптимизационни процедури са изпълнени при четири различни начални стойности на параметрите. Първата и втора група начални стойности са съответните долна и горна граница на изменение за всеки един параметър, съгласно Таблица 1. Третата група е получена при усредняване на допустимите граници на изменение на параметрите, съгласно Таблица 1, а началните стойности в четвъртата група са избрани на случаен принцип, като е спазено единствено условието да бъдат в границите посочени в Таблица 1.

За сравнение на четирите традиционни метода са използвани два критерия: време за намиране на решението и стойност на целевата функция. Най-добрите резултати с всеки един от приложените методи са представени в Таблица 2.

Въз основа на представените в Таблица 2 резултати може да се обобщи, че получените при приложение на метода на *Гаус-Нютон* и *Левенберг-Маркуардт* оценки, дублират зададените начални стойности на съответния параметър и поради тази причина не могат да бъдат приети за надеждни. Съгласно постигнатата стойност на оптимизационния критерий, методът на деформируемия симплекс дава по-добри резултати от минимаксния метод. Но също така е очевидно, че пет от общо деветте оценявани параметъра, а именно μ_{2S} , k_S , Y_{SX} , Y_{EX} и Y_{OS} са извън посочените в Таблица 1. граници. От Таблица 2 се вижда още, че конвенционалните оптимизационни алгоритми достигат до различни решения - вероятно локални

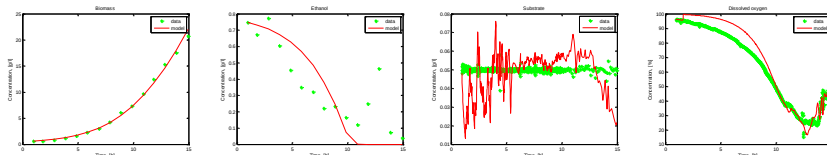
минимуми. Това може да се дължи и на факта, че изследваният модел е сложен и нелинеен, което предполагаемо води до многоекстремалност на оптимизационния критерий J .

Таблица 2 Резултати от параметрична идентификация с класически методи за оптимизация

	Метод на деформируеия симплекс	Минимаксен метод	Метод на Левенберг-Маркуардт	Метод на Гаус-Нютон
J	0,0245	0,0592	0,0633	0,0633
t , s	95,8430	54,187	3,75	3,734
μ_{2S} , 1/h	1,0449	0,9	0,53	0,53
μ_{2E} , 1/h	0,1554	0,3	0,1	0,1
k_S , g/l	0,0886	0,5	0,12	0,12
k_E , g/l	0,7238	4,9	0,5	0,5
Y_{SX} , g/g	0,81	0,05	0,37	0,37
Y_{EX} , g/g	2,26	0,17	0,24	0,24
$k_L^{O_2} a$, 1/h	444,7658	1199,9	420	420
Y_{OS} , g/g	311,3358	1199	450	450
Y_{OE} , g/g	0,5202	8340	0,9	0,9

Като следващ етап от изследването е извършена постъпкова параметрична идентификация [29] на модела на полупериодичната култивация на дрожди *S. cerevisiae*, с метода на деформируеия симплекс. За разглеждания набор от експериментални данни, описван с помощта на модел (1)-(5) и специфични скорости (6) постъпковата параметрична идентификация се изразява в разделянето на деветте оценявани параметъра на четири групи като се стартира с идентификация по субстрат и набор от параметри на модела $p_1 = [\mu_{2S}, k_S, Y_{SX}]$. След това се извършва идентификация по биомаса при набор от параметри на модела $p_2 = [\mu_{2E}, k_E]$, продължава се с идентификация по етанол с набор от параметри на модела $p_3 = [Y_{EX}]$ и се приключва с идентификация по разтворен кислород със съответен набор от параметри $p_4 = [k_L^{O_2} a, Y_{OS}, Y_{OE}]$. На всяка следваща стъпка от процедурата се използват оценките на параметрите, идентифицирани на предходен етап.

Усреднените стойности на допустимите граници на изменение на параметрите, съгласно Таблица 1 са използвани като начални, тъй като при тях методът на деформируеия симплекс достига до най-ниска стойност на оптимизационния критерий. В описаната по-горе последователност, началните стойности на параметрите са както следва: $p_{1,0} = [0.56, 0.3, 0.35]$, $p_{2,0} = [0.18, 2.56]$, $p_{3,0} = [1.12]$ и $p_{4,0} = [628, 799, 4170]$, докато оценените стойности на параметрите, в същата последователност, са съответно $p_{1,ident} = [0.6331, 0.3210, 0.2708]$, $p_{2,ident} = [0.2205, 0.0030]$, $p_{3,ident} = [8.4735]$, $p_{4,ident} = [0.1843e^3, 2.8186e^3, -0.5894e^3]$. Сумарният оптимизационен критерий от идентификациите по субстрат, биомаса, етанол и разтворен кислород е $J = 0,0301$ докато сумарното време, в секунди, за намиране на решение е $t = 970,5781$. На Фиг. 1 е представен графично резултатът след постъпковата параметрична идентификация на модела с приложение на метода на деформируеия симплекс.



Фиг. 1 Експериментални и моделни данни за концентрациите на биомаса, етанол, субстрат и разтворен кислород при постъпкова идентификация на параметри с метод на деформируемия симплекс

Въпреки постигнатата точност на припокриване между експерименталните и моделно получените данни, отрицателната оценка на параметъра Y_{OE} , от биологична гледна точка, е лишена от смисъл, което отново показва невъзможността на традиционните методи за оптимизация да се справят адекватно със задачата за параметрична идентификация на изследвания модел на ферментационен процес.

Параметрична идентификация с еднopoпулационни генетични алгоритми

Както вече беше показано, традиционните оптимизационни методи не успяват да преодолеят спецификата на ФП и не достигат до желано решение на следващ етап, като алтернатива за параметричната идентификация на изследвания процес, са тествани възможностите на генетичните алгоритми. Съгласно [15, 16], принципът на работа на СГА може да бъде описан в осем стъпки:

1. **[Старт]**
Генериране на случайна популация с n на брой хромозоми
2. **[Целева функция]**
Изчисляване на целевата функция за всяка хромозома n в популацията
3. **[Фитнес функция]**
Изчисляване на фитнес функцията за всяка хромозома n в популацията
4. **[Нова популация]**
Създаване на нова популация, чрез повтаряне на следните стъпки:
 - 4.1. **[Селекция]**
Избор на родителски хромозоми от популацията съгласно техните фитнес функции
 - 4.2. **[Кръстосване]**
Кръстосване, с определена вероятност, на родителските хромозоми, с цел формиране на ново поколение
 - 4.3. **[Мутация]**
Мутиране, с определена вероятност, на новото поколение
5. **[Приемане]**
Поставяне на новото поколение в новата популация
6. **[Заместване]**
Използване на новогенерираната популация за бъдещо стартиране на алгоритъма

7. **[Тест]**

Ако условието за край е изпълнено, алгоритъмът спира и връща най-доброто намерено решение в текущата популация, ако условието за спиране на алгоритъма не е изпълнено, алгоритъмът продължава

8. **[Цикъл]**

Връщане към стъпка **Фитнес функция**.

Разработване на модификации на СГА

Както беше споменато вече в стандартния еднопопулационен генетичен алгоритъм [15] хромозомите (кодиран набор от параметри), представляващи възможни решения на проблема, се избират от поколението според стойностите на собствените им обектни функции, посредством генетичен оператор *селекция*. Така избраните хромозоми се подлагат на *кръстосване*, за да се формира ново поколение. Новото поколение може да претърпи *мутация* с определена вероятност, за разнообразяване на популацията и предотвратяване попадането на решенията в локален екстремум. В настоящата дисертация този алгоритъм ще бъде означен като СГА_СКМ.

Правейки аналогия с протичащите в природата процеси може да се каже, че вероятността да се осъществи първо мутация в индивидите, след което те да бъдат кръстосани е съпоставима с вероятността двата процеса да се случват в обратна последователност. Следвайки тази логика може да се твърди също, че селекцията на индивидите е възможно да бъде извършена преди, след или между процесите кръстосване и мутация, както е възможно, например, индивидите да бъдат подложени само на кръстосване и селекция, без да мутират. Следвайки гореописаната логика общо пет нови модификации (тук означени съответно като СГА_СМК, СГА_МКС, СГА_КСМ, СГА_МСК и СГА_КС) на три известни от литературата СГА (тук означени съответно като СГА_СКМ [15], СГА_КМС [28], СГА_СК [2, 30]) са разработени в настоящата дисертация. За изследваните общо осем вида СГА последователността на изпълнение на осемте стъпки, представени по-горе е както следва:

- **СГА_СКМ** 1, 2, 3, **4.1, 4.2, 4.3**, 5, 6, 7, 8
- **СГА_КМС** 1, 2, 3, **4.2, 4.3, 4.1**, 5, 6, 7, 8
- **СГА_СМК** 1, 2, 3, **4.1, 4.3, 4.2**, 5, 6, 7, 8
- **СГА_МКС** 1, 2, 3, **4.3, 4.2, 4.1**, 5, 6, 7, 8
- **СГА_КСМ** 1, 2, 3, **4.2, 4.1, 4.3**, 5, 6, 7, 8
- **СГА_МСК** 1, 2, 3, **4.3, 4.1, 4.2**, 5, 6, 7, 8
- **СГА_СК** 1, 2, 3, **4.1, 4.2**, 5, 6, 7, 8
- **СГА_КС** 1, 2, 3, **4.2, 4.1**, 5, 6, 7, 8

Изследване влиянието на параметрите на СГА

Параметричната идентификация на модела (1)-(5) и специфични скорости (6) е проведена в среда *Matlab 7* с използването на *Genetic Algorithm Toolbox* [13, 14]. За числената симулация на моделите е приложен интеграционният алгоритъм *Runge-Kutta (RK45)*. Скаларът на отклонението от относителната грешка *RelTol* е зададен като $1e^{-4}$, докато векторът на отклонение от абсолютната грешка (за всички параметри) *AbsTol* е

прецизиран на $1e^{-5}$. Всички изчисления са извършени на *PC Intel Pentium 4 (2,4 GHz)* с платформа *Windows XP*.

В Таблица 3 са представени стойностите, за които са изследвани параметрите на ГА, съгласно някои твърдения в [22]:

Таблица 3. Стойности на изследваните параметри на ГА

NIND	MAXGEN	GGAP	XOVR	MUTR
20	100	0.5	0.65	0.02
40	200	0.67	0.75	0.04
60	500	0.8	0.85	0.06
80	1000	0.9	0.95	0.08
100	-	-	-	0.1

Анализът на ефекта на параметрите на ГА е извършен поотделно за всеки един параметър, като изследването на всяка стойност за съответния параметър става при константни стойности на останалите параметри. Когато един от петте параметъра на ГА – NIND, MAXGEN, GGAP, XOVR или MUTR се изследва за стойностите, показани в Таблица 3, стойностите на другите параметри, съгласно [22] са установени на: NIND = 20, MAXGEN = 100, GGAP = 0.8, XOVR = 0.95 и MUTR = 0.05.

Стойностите на останалите параметри и вида на генетичните оператори, използвани при различните СГА, тествани за параметричната идентификация на модела на полупериодичен процес за култивиране на дрожди *S. cerevisiae* са представени съответно в Таблица 4 и Таблица 5.

Таблица 4. Параметри на СГА

Параметър	Стойност
NVAR	9
PREC1	20

където: NVAR – брой на променливите; PREC1 – точност на бинарното представяне.

Таблица 5. Вид на използваните оператори в СГА

Оператор	Вид
Кодиране	Двоично
Повторно въвеждане	Базирано на стойността на фитнес функцията
Кръстосване	Двучовково
Мутация	Битова инверсия
Селекция	Метод на рулетката
Фитнес функция	Линеен ранкинг

Поради стохастичния характер на ГА всички анализи в дисертацията са извършени на базата на получените усреднени стойности за целевата функция, времето за сходимост и стойностите на параметрите на модела. За тази цел всички изследвани алгоритми са изпълнявани по тридесет пъти, за всеки конкретен случай, след което резултатите са усреднени и представени съответно в табличен вид.

За по-лесно сравняване на получените резултати, осемте вида СГА могат да бъдат групирани по два. Три от групите са формирани в зависимост от това дали операторът *селекция* се изпълнява преди, след или между операторите *кръстосване* и *мутация* (СГА_СКМ и СГА_СМК; СГА_КМС и СГА_МКС; СГА_КСМ и СГА_МСК), а четвъртата се състои от двата алгоритъма без изпълнение на оператора мутация (СГА_СК и СГА_КС).

В Таблица 6 са представени резултатите от параметричната идентификация на полупериодична култивация на дрожди *S. cerevisiae* със СГА_СКМ, СГА_СМК, СГА_КМС и СГА_МКС при различни стойности на GGAP, тъй като този параметър на ГА е отличен като най-чувствителен по отношение на бързодействието на изследваните алгоритми.

Таблица 6. Резултати от параметричната идентификация с различни видове СГА

GGAP	СГА СКМ		СГА СМК		СГА КМС		СГА МКС	
	J	t, s	J	t, s	J	t, s	J	t, s
0.5	0.0223	43.8120	0.0221	44.5160	0.0222	38.1710	0.0221	37.2810
0.67	0.0221	52.7970	0.0221	55.0460	0.0225	37.1870	0.0223	39.7810
0.8	0.0221	67.9220	0.0221	67.6410	0.0222	50.2970	0.0224	53.4530
0.9	0.0222	70.6250	0.0222	70.9690	0.0224	52.3440	0.0230	61.4370

Както е показано в Таблица 6, стойностите на оптимизационния критерий, получени с четирите вида стандартни генетични алгоритми, са много сходни. Те варират между 0.0221 и 0.0230, което означава, че се наблюдава само около 4% разлика в точността на модела. Следва да се отбележи, че измежду всички изследвани параметри GGAP е най-чувствителният по отношение на времето за сходимост на СГА. В случая на СГА_МКС, отличен и като най-бързия алгоритъм в разглежданата група от алгоритми, когато се използва стойност на GGAP = 0.5 вместо 0.9 времето за намиране на решение може да бъде намалено с до почти 40% без загуба на точност на модела. Дори нещо повече, точно в случая на СГА_МКС стойността на целевата функция намалява от най-високата 0.0230, измежду всички наблюдавани стойности, до най-ниската – 0.0221.

Въпреки, че не са показани, резултатите за останалите изследвани параметри на ГА са подробно анализирани. И при четирите вида СГА стойността на целевата функция намалява твърде нечувствително с 0.01% когато стойността на MAXGEN се увеличава, но за сметка на изчислителното време на алгоритъма. Поради тази причина стойност за MAXGEN = 100 е приета като добър компромис между броя изчисления на целевата функция, точността на модела и времето за намиране на решение. За другия параметър на ГА - NIND, отново изследван и при четирите вида СГА, стойността на целевата функция расте с 0.01%, когато стойността на броя индивиди в популацията се увеличава. Времето за намиране на решение

също расте. Поради тази причина стойност за $NIND = 20$ е възприета като подходящ компромис между точността на решението и времето, необходимо на алгоритъма да достигне до оптимално решение. Въпреки, че изследването на различни стойности на $XOVR$ не води до намаляване на времето за намиране на глобален минимум, може да се каже, че стойност 0.85 за $XOVR$ се препоръчва като по-подходяща. Що се отнася до параметъра $MUTR$, при него няма ясно открояваща тенденция за влиянието му върху бързодействието на алгоритмите и точността на намереното решение. За фаворита измежду разглежданите четири алгоритъма CGA_MKS , стойност за $MUTR = 0.1$ може да спести до 20% от времето за сходимост на алгоритъма.

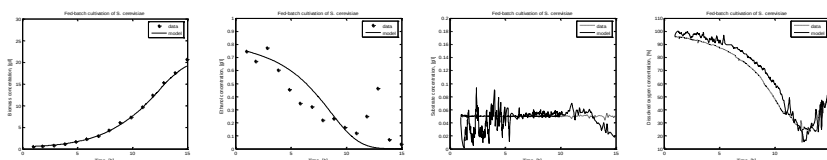
Сходните стойности на оптимизационния критерий, получени при приложението на четирите вида CGA гарантира, че от гледна точка на точност изпълнението на оператора *мутация* преди оператора *кръстосване* не води до намаляване степента на адекватност на модела. Още повече, че предложените модификации в CGA в повечето случаи намаляват времето за достигане до глобален минимум. Представеното сравнение показва, че изпълнението на операторите в последователност *мутация*, *кръстосване* и *селекция* е оптимално по отношение на времето за сходимост на алгоритъма, с гарантирана висока адекватност на намереното решение. Въз основа на този анализ, като „фаворит“ измежду изследваните алгоритми, при които операторът *селекция* се изпълнява съответно преди или след операторите *кръстосване* и *мутация*, може да бъде отличен CGA_MKS .

С цел демонстрация на получените резултати CGA_MKS е приложен за параметрична идентификация на полупериодичната култивация на дрожди *S. cerevisiae*, като са използвани избраните стойности на параметрите на GA : $NIND = 20$, $MAXGEN = 100$, $GGAP = 0.5$, $XOVR = 0.85$ и $MUTR = 0.1$. В Таблица 7 са представени оценените стойности на параметрите на модела, като изчислителното време на алгоритъма е $CPU\ time = 38.6410$, а стойността на целевата функция е $J = 0.0223$.

Таблица 7 Оценени стойности на параметрите на модела с приложение на CGA_MKS

Параметър	μ_{2S} [1/h]	μ_{2E} [1/h]	k_S [g/l]	k_E [g/l]	γ_{SX} [g/g]	γ_{EX} [g/g]	k_{La} [1/h]	γ_{OS} [g/g]	γ_{OE} [g/g]
Стойност	0.95	0.12	0.12	0.80	0.4	1.47	113.02	894.88	254.95

На Фиг. 2 са представени резултатите от експерименталните и моделно предсказани данни със CGA_MKS , респективно за концентрациите на биомасата, етанола, субстрата и разтворения кислород.



Фиг. 2 Експериментални и моделно предсказани данни със CGA_MKS за концентрациите на биомаса, етанол, субстрат и разтворен кислород

Влиянието на петте параметъра NIND, MAXGEN, GGAP, XOVR и MUTR върху времето за сходимост на алгоритмите и степента на точност на модела е изследвано и за втората група от четири СГА. Два от тях СГА_KCM и СГА_MCK са новоразработени въз основа на идеята операторът *селекция* да се изпълнява между операторите *кръстосване* и *мутация*, а другите два алгоритъма - СГА_CK и СГА_KC си приличат по това, че при изпълнението им не се включва операторът *мутация*, като СГА_KC е също новопредложен. Получените резултати при изследване влиянието на NIND, MAXGEN, XOVR и MUTR бяха подробно анализирани, но не са представени тъй като както и при първата група от четири алгоритъма не водят до чувствителна промяна във времето за сходимост на алгоритъма и степента на точност на модела. Следва да се отбележи, че GGAP е отново най-чувствителният измежду изследваните пет параметъра по отношение на времето за сходимост. Ето защо в Таблица 8 са представени резултатите от параметрична идентификация на разглежданата полупериодична култивация на дрожди *S. cerevisiae*, при приложение на новите четири вида СГА само за стойностите на отличения вече като най-чувствителен параметър по отношение времето за намиране на решение - GGAP.

Таблица 8. Резултати от параметричната идентификация с различни видове СГА

GGAP	СГА_KCM		СГА_MCK		СГА_KC		СГА_CK	
	J	t, s	J	t, s	J	t, s	J	t, s
0.5	0.0224	36.281	0.0223	36.266	0.0230	31.593	0.0223	30.735
0.67	0.0223	39.203	0.0228	53.547	0.0248	41.375	0.0235	43.953
0.8	0.0223	60.313	0.0223	47.485	0.0233	41.250	0.0228	42.578
0.9	0.0235	55.921	0.0223	77.609	0.0230	53.734	0.0237	39.359

И тук се очертава наблюдаваната при предходните четири алгоритъма тенденция, че сходните структури на СГА водят до получаване на съпоставими резултати по отношение на стойността на целевата функция и времето за сходимост на алгоритмите. СГА_KCM и СГА_MCK са по-точни в намереното решение, съгласно постигнатите стойности на целевата функция, но са и по-бавни в сравнение със СГА_KC и СГА_CK. Двата алгоритъма без изпълнение на оператора *мутация* - СГА_KC и СГА_CK могат да бъдат групирани заедно като най-бързи. Съвсем логично е в този случай, когато ГА се изпълняват само с два оператора (*селекция* и *кръстосване*), изчислителното време на алгоритъма да намалее, но в повечето случаи за сметка на точността на намереното решение. В случая на СГА_CK, който е отличен и като най-бързия измежду така групирани и изследвани алгоритми до 28% от времето за намиране на решение могат да бъдат спестени когато се използва GGAP = 0.5 вместо 0.9, при това с чувствително повишаване на степента на точност.

За останалите четири параметъра са получени аналогични резултати с тези от предходната група изследвани СГА. Затова стойности на параметрите NIND = 20, MAXGEN = 100, GGAP = 0.5, XOVR = 0.85 и MUTR = 0.1 са препоръчани като най-подходящи.

При сравнение на резултатите за всички общо осем изследвани СГА, логично най-бързи се оказват двата алгоритъма без изпълнение на оператора *мутация*, следвани от двата новоразработени алгоритъма, при които операторът *селекция* се изпълнява между другите два оператора за репродукция на индивидите.

Параметрична идентификация с многопопулационни генетични алгоритми

Механизмът на изпълнение на МПГА е по-близък до процесите протичащи в живата природа. Следователно неговото приложение за целите на параметрична идентификация на изследвания модел на полупериодична култивация на дрожди *S. cerevisiae* е обосновано от логично очакваната възможност за подобряване на получените със СГА резултати.

Принципът на работа на стандартния МПГА, съгласно [15, 16] е описан в девет стъпки:

1. **[Старт]**
Генериране на k случайни субпопулации всяка, от които с n на брой хромозоми
2. **[Целева функция]**
Изчисляване на целевата функция за всяка хромозома n в субпопулациите
3. **[Фитнес функция]**
Изчисляване на фитнес функцията за всяка хромозома n в субпопулациите
4. **[Нова популация]**
Създаване на нови субпопулации чрез повтаряне на следните стъпки:
 - 4.1. **[Селекция]**
Избор на родителски хромозоми от субпопулациите, съгласно техните фитнес функции
 - 4.2. **[Кръстосване]**
Кръстосване, с определена вероятност, на родителските хромозоми, с цел формиране на поколения
 - 4.3. **[Мутация]**
Мутиране, с определена вероятност, на новите поколения
5. **[Приемане]**
Поставяне на новите поколения в новите субпопулации
6. **[Заместване]**
Използване на новогенерираните субпопулации за последващо стартиране на алгоритъма
7. **[Миграция]**
Мигриране на определен брой индивиди между субпопулациите след изтичане на времето за изолация
8. **[Тест]**
Ако условието за край е изпълнено, алгоритъмът спира и връща най-доброто намерено решение, ако условието за спиране на алгоритъма не е изпълнено, алгоритъмът продължава.
9. **[Цикъл]**
Връщане към стъпка **Фитнес функция**.

Разработване на модификации на МПГА

Отново следвайки еволюционните механизми, възприети като идея при разработването на СГА, общо шест нови модификации (тук означени съответно като МПГА_КМС, МПГА_СМК, МПГА_МКС, МПГА_КСМ, МПГА_МСК и МПГА_КС) на два известни от литературата МПГА (тук означени съответно като МПГА_СКМ [15], МПГА_СК [2]) са предложени и изследвани за целите на параметричната идентификация на полупериодична култивация на дрождите *S. cerevisiae*.

За изследваните осем МПГА, последователността на изпълнение на деветте стъпки, описани по-горе е както следва:

- **МПГА_СКМ** 1, 2, 3, **4.1**, **4.2**, **4.3**, 5, 6, 7, 8, 9
- **МПГА_КМС** 1, 2, 3, **4.2**, **4.3**, **4.1**, 5, 6, 7, 8, 9
- **МПГА_СМК** 1, 2, 3, **4.1**, **4.3**, **4.2**, 5, 6, 7, 8, 9
- **МПГА_МКС** 1, 2, 3, **4.3**, **4.2**, **4.1**, 5, 6, 7, 8, 9
- **МПГА_КСМ** 1, 2, 3, **4.2**, **4.1**, **4.3**, 5, 6, 7, 8, 9
- **МПГА_МСК** 1, 2, 3, **4.3**, **4.1**, **4.2**, 5, 6, 7, 8, 9
- **МПГА_СК** 1, 2, 3, **4.1**, **4.2**, 5, 6, 7, 8, 9
- **МПГА_КС** 1, 2, 3, **4.2**, **4.1**, 5, 6, 7, 8, 9

Изследване влиянието на параметрите на МПГА

Влиянието на седем основни параметъра на МПГА е изследвано съобразно точността на намереното решение и времето за сходимост на алгоритмите. Пет от параметрите са идентични по вид и стойност с разгледаните вече при СГА (Таблица 3), а останалите два са съответно: вероятност за включване (INSR), изследван за стойности {0.5, 0.6, 0.8, 0.9, 1} и вероятност за миграция (MIGR), изследван за стойности {0.1, 0.2, 0.4, 0.6, 0.8}.

При изследване ефекта на всеки един параметър, стойностите на останалите параметри са установени на: NIND = 20, MAXGEN = 100, GGAP = 0.8, XOVR = 0.95, MUTR = 0.05, INSR = 0.95 и MIGR = 0.2.

Стойностите на неизследваните параметри при МПГА и вида на операторите са идентични с използваните при СГА (Таблицы 4 и 5). В Таблица 9 са дадени стойностите на параметрите, допълнително необходими за изпълнение само на МПГА и съответно невключени в Таблица 4.

Таблица 9. Параметри на МПГА

Параметър	Стойност
SUBPOP	5
MIGGEN	20

където SUBPOP – брой субпопулации; MIGGEN – брой генерации, след които се осъществява миграция между субпопулациите.

Влиянието на параметрите, а именно NIND, MAXGEN, GGAP, XOVR, MUTR, INSR и MIGR е изследвано за шестте вида МПГА, в които се изпълняват трита оператора *селекция*, *кръстосване*, *мутация*. Параметърът MUTR не е изследван само при МПГА_СК и МПГА_КС, поради липсата на оператор *мутация* в тези два алгоритъма. Параметрите са изследвани един по един последователно за всяка от избраните стойности. Бързодействието

на всеки от изследваните алгоритми и стойността на целевата функция са представителните критерии, използвани за това сравнение.

В Таблица 10 са представени резултатите от изследването на GGAP за всички осем вида МПГА. Получените резултати при изследване влиянието на MAXGEN, NIND, HOVR, MUTR, INSR и MUTR бяха подробно анализирани, но не са представени тъй като не водят до чувствителна промяна във времето за сходимост на алгоритъма и степента на точност на модела.

Таблица 10. Влияние на GGAP върху времето за сходимост и стойността на целевата функция за осемте вида МПГА

GGAP	МПГА СКМ		МПГА СМК		МПГА КМС		МПГА МКС	
	J	t, s	J	t, s	J	t, s	J	t, s
0.5	0.0220	100.8910	0.0220	111.7810	0.0221	273.9060	0.0220	307.8440
0.67	0.0221	112.1720	0.0220	141.0940	0.0221	325.5780	0.0220	332.0620
0.8	0.0221	155.4680	0.0220	178.9680	0.0221	321.0160	0.0221	373.1560
0.9	0.0220	170.2660	0.0220	340.6720	0.0221	343.6870	0.0221	349.7500

GGAP	МПГА СК		МПГА КС		МПГА КСМ		МПГА МСК	
	J	t, s	J	t, s	J	t, s	J	t, s
0.5	0.0222	111.5310	0.0223	267.9220	0.0221	97.6870	0.0220	98.2970
0.67	0.0224	119.7340	0.0222	331.9690	0.0221	128.8750	0.0221	121.8600
0.8	0.0221	153.3900	0.0223	333.6250	0.0221	163.8590	0.0220	145.6710
0.9	0.0220	168.2190	0.0221	357.0160	0.0221	165.6720	0.0220	166.0150

Може да се отбележи, че резултатите, получени с МПГА_СКМ са много сходни с резултатите, получени с МПГА_СМК и са съпоставими с резултатите от другата двойка модификации МПГА_КСМ и МПГА_МСК. Също така, резултатите, когато се прилага МПГА_КМС са близки до тези когато е използван МПГА_МКС, но времето за сходимост при тази двойка МПГА е много по-голямо от това на първите две двойки. Вижда се също, че няма загуба на точност на модела, когато операторът *мутация* се изпълнява преди оператора *кръстосване*. От всичко гореизложено може да се обобщи, че изпълнението на оператора *селекция* преди и между операторите *кръстосване* и *мутация* (без значение от техния ред) се нуждае от много по-малко изчислителни време. Този факт е валиден за всички изследвани МПГА и за всички изследвани параметри на ГА. Двата алгоритъма без изпълнение на оператора *мутация*, МПГА_СК и МПГА_КС, също могат да бъдат групирани заедно. В случаите, когато алгоритмите се изпълняват само с два оператора (*селекция* и *кръстосване*) точността логично намалява, както беше установено и при СГА.

Следва да се отбележи, че GGAP отново е най-чувствителният от седемте изследвани параметъра на ГА по отношение времето за сходимост на алгоритъма (Таблица 10). До 41% от изчислителното време на алгоритъма (в случая на МПГА_КСМ, който е отличен и като най-бързия алгоритъм при изследване на GGAP) могат да бъдат спестени без загуба на точност, ако се използва GGAP = 0.5 вместо 0.9. Но ако бъдат сравнени МПГА_КСМ с GGAP = 0.5 (най-бърз) и МПГА_МКС с GGAP = 0.8 (най-бавния), то първият алгоритъм е 3.8 пъти по-бърз от втория.

След анализ на резултатите за осемте вида МПГА стойност за MAXGEN = 100 е приета като добър компромис между броя изчисления, точността на решението и времето за намиране на решение. За изследвания параметър NIND, стойност за NIND = 20 е възприета като подходящ компромис между точността на модела и изчислителното време на алгоритъма.

Изследването на различните стойности за параметъра вероятност за кръстосване не води до намаляване на времето за сходимост, но следва да се отбележи, че стойност 0.85 за XOVR може да се приеме като най-подходяща за всички разгледани МПГА алгоритми. Всичките шест алгоритъма, при които се изпълнява операторът мутация, са най-бързи при MUTR = 0.02. Ето защо MUTR = 0.02 може да се приеме като най-подходяща, спестявайки до 24% от изчислителното време на алгоритъма в случая на МПГА_MCK, без това да води до загуба на точност на модела. Въз основа на получените резултати тенденция на влияние на генетичните параметри INSR и MIGR не е намерена.

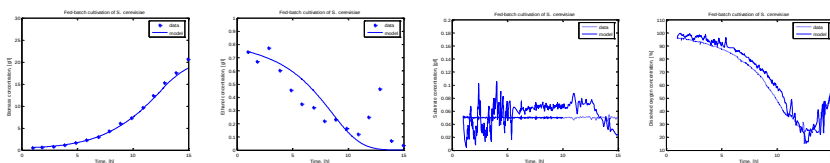
Сравнението на осемте алгоритъма показва, че за един от изследваните параметри (MIGR) на ГА МПГА_CKM е най-бързият алгоритъм, а за останалите четири параметъра стандартният МПГА отстъпва първенството на някоя от новопредложените тук модификации. При изследване на параметрите GGAP и INSR най-бърз се оказва новоразработеният МПГА_KCM, докато при XOVR и MUTR най-бързи съответно са МПГА_CMK и МПГА_MCK.

Поради сходство на получените резултати, тук е представена параметричната идентификация на модел на полупериодичен процес на дрожди *S. cerevisiae* с една от новопредложените модификации, а именно МПГА_KCM, като са използвани предложените стойности на параметрите на ГА: NIND = 20, MAXGEN = 100, GGAP = 0.5, XOVR = 0.85, MUTR = 0.02, INSR = 0.9 и MIGR = 0.1. В Таблица 11 са представени оценените стойности на параметрите на модела, като изчислителното време на алгоритъма е CPU time = 97.5940, а стойността на целевата функция е $J = 0.0221$.

Таблица 11 Оценени стойности на параметрите на модела с приложение на МПГА_KCM

Параметър	μ_{2S} [1/h]	μ_{2E} [1/h]	k_S [g/l]	k_E [g/l]	Y_{SX} [g/g]	Y_{EX} [g/g]	$k_L a$ [1/h]	Y_{OS} [g/g]	Y_{OE} [g/g]
Стойност	0.90	0.12	0.15	0.80	0.41	1.64	65.20	509.82	360.17

На Фиг. 3 са показани резултатите от експерименталните и моделно предсказани данни за концентрациите на биомаса, етанол, субстрат и разтворен кислород.



Фиг. 3 Експериментални и моделно предсказани данни с МПГА_KCM за концентрациите на биомаса, етанол, субстрат и разтворен кислород

ИЗВОДИ

За целите на параметричната идентификация на модел на полупериодична култивация на дрожди *S. cerevisiae* са изследвани възможностите на четири традиционни метода за оптимизация – метод на деформируемия *симплекс*, *минимаксен* метод, метод на *Маркуардт-Левенбег* и метод на *Гаус-Нютон*. Най-точно решение, съгласно достигнатата стойност на целевата функция, дава методът на деформируемия *симплекс*, но въпреки това резултатите демонстрират невъзможността на традиционните методи за оптимизация да отразят специфичните особености на моделите на ФП. На следващ етап, като най-точен, методът на деформируемия *симплекс* е избран за извършване на постъпкова параметрична идентификация на модела. Процедурата се прилага, за да бъдат подобрени получените до момента резултати, но този път отрицателната оценка за един от параметрите обезсмисля от биологична гледна точка намереното решение. Ето защо възниква идеята за параметричната идентификация на разглеждания модел на ФП да бъдат тествани генетичните алгоритми, като стохастичен метод с широко приложение за решаването на различни сложни задачи.

В настоящата глава от дисертацията за целите на параметричната идентификация на полупериодичен процес за култивиране на дрожди *S. cerevisiae* са изследвани възможностите на генетичните алгоритми като стохастична техника за намиране на глобален екстремум. Изследвани са общо шестнадесет вида ГА - осем вида еднopoпулационни и осем вида многопопулационни ГА. Различните модификации на СГА и МПГА са получени след размяна на последователността на изпълнение на основните генетични оператори *селекция*, *кръстосване* и *мутация*, а четири от представените ГА се изпълняват без оператора *мутация*.

Влиянието на по-важните параметри на ГА, а именно брой на индивидите в популацията (NIND), брой генерации (MAXGEN), степен на прехвърляне (GGAP) и вероятност за кръстосване (XOVR) е изследвано за всички шестнадесет вида ГА, с цел подобряване стойността на намереното решение и времето за сходимост на алгоритмите. Влиянието на параметъра вероятност за мутация (MUTR) е изследван за шест вида СГА и шест вида МПГА, при които се изпълнява операторът мутация. За осемте МПГА е изследвано и влиянието на още два параметъра, специфични само за този вид ГА, а именно вероятност за вмъкване (INSR) и вероятност за миграция (MIGR). Измежду изследваните параметри на ГА, GGAP се оказва най-чувствителен по отношение на времето за достигане до решение, както при осемте вида СГА така и при осемте вида МПГА. Съобразно проведеното изследване като „фаворити“ измежду разглежданите с три оператора СГА, могат да бъдат отличени СГА_MKC, СГА_KCM, СГА_MCK, а измежду разглежданите МПГА - МПГА_KCM, МПГА_CMK и МПГА_MCK. Съответно до почти 40% и 41% от изчислителното време на алгоритъма може да бъде спестено в случаите на СГА_MKC и МПГА_KCM, при които се използва GGAP = 0.5 вместо 0.9, без това да води до загуба на точност на модела.

При изследване на различните стойности за другите параметри не се реализира същото намаляване на времето за намиране на решение, но следва да се отбележи, че стойностите NIND = 20, MAXGEN = 100,

XOVR = 0.85, MUTR = 0.1 (0.02 за МПГА) и MIGR = 0.1, INSR = 0.9 се приемат като най-подходящи.

Използването на избраните стойности на параметрите и предложените модификации, съответно при СГА и МПГА, подобрява ефективността на разглежданите алгоритми, по отношение на тяхната сходимост, без загуба на точност на модела. Това обуславя целесъобразното им прилагане за решаване на комплексни нелинейни задачи, каквито са задачите за параметрична идентификация на модели на ФП.

Глава 3

Целенасочен параметричен генезис при параметрична идентификация на модел на полупериодична култивация на дрожди *S. cerevisiae*

Поради стохастичния характер на ГА са необходими многократни изпълнения на алгоритъма, които да гарантират получаване на надеждни резултати. Първоначално генетичният алгоритъм търси стойностите на параметрите на модела в широки, но разумно избрани граници. При анализ на натрупаните бази данни от многократни изпълнения на ГА резултатите показват, че е налице групиране на стойностите на параметрите в по-тесни граници. Това провокира идеята за разработване на процедура за целенасочен генезис на параметрите на модела чрез стесняване на предварително определените граници, с цел да се намали времето за намиране на решение, без това да се отрази върху точността на модела. Процедурата за целенасочен параметричен генезис е разработена първоначално за еднопопулационен и след това приложена и за многопопулационен ГА.

Процедурата за целенасочен параметричен генезис се представя в шест последователни стъпки:

Стъпка 1: Извършват се N на брой изпълнения на генетичния алгоритъм.

Стъпка 2: Определя се минимална (min) и максимална (max) стойност на целевата функция.

Стъпка 3: Определя се високо ниво (ВН), средно ниво (СН) и ниско ниво (НН) на представяне на алгоритъма със съответните долна (ДГ) и горна (ГГ) граница, съгласно схемата:

3.1. Определяне на дискриминационната стойност Δ чрез

$$\Delta = \frac{\max J - \min J}{3}$$

3.2.	Високо ниво долна граница (ВН_ДГ)	$\min J$
	Високо ниво горна граница (ВН_ГГ)	$\min J + \Delta - \varepsilon$
	Средно ниво долна граница (СН_ДГ)	$\min J + \Delta$
	Средно ниво горна граница (СН_ГГ)	$\min J + 2\Delta - \varepsilon$
	Ниско ниво долна граница (НН_ДГ)	$\min J + 2\Delta$
	Ниско ниво горна граница (НН_ГГ)	$\max J$

където ε е число, което осигурява разграничаването на

нивата.

Стъпка 4: Определяне на минимална, средна и максимална стойност за всеки параметър във всяко от горепосочените нива.

Стъпка 5: Определяне на новите интервали на изменение на параметрите на модела, на базата на средните стойности във високото ниво.

Стъпка 6: Изпълнение на генетичния алгоритъм с интервалите, определени в Стъпка 5.

Постъпковата процедура преминава през всички описани по-горе стъпки, без да се пропуска нито една от тях и без да се изпълняват цикли. Най-надеждните решения по отношение на целевата функция попадат в интервала $[min J; min J + \Delta - \epsilon]$. Решенията от средното ниво на представяне имат стойност на целевата функция изменяща се в интервала $[min J + \Delta; min J + 2\Delta - \epsilon]$. Най-ненадеждните решения попадат в интервала $[min J + 2\Delta; max J]$.

Приложение на процедура за целенасочен параметричен генезис при различни видове СГА

Приложение на процедура за целенасочен параметричен генезис при стандартен СГА

Разработената процедура за целенасочен параметричен генезис е приложена за параметрична идентификация на полупериодична култивация на дрожди *S. cerevisiae*. Общо деветте параметъра на модела (1)-(5) със специфични скорости (6) са оценени след прилагане на стандартен еднопопулационен генетичен алгоритъм (СГА_СКМ), с последователност на изпълнение на основните генетични оператори *селекция*, *кръстосване*, *мутация*. СГА_СКМ е изследван за четири различни стойности на най-чувствителния параметър по отношение на времето за сходимост, а именно GGAP. СГА_СКМ е изпълнен по тридесет пъти за всяка стойност на GGAP. Получените резултати са анализирани в зависимост от постигнатите стойности на целевата функция и са представени в Таблица 12.

Във всяко едно от нивата на представяне на алгоритмите, построени съгласно ПЦПГ, се определя минимумът, максимумът и средната стойност за всеки един параметър от модела. Таблица 13 представя тези стойности, след закръгляване до втори знак след десетичната запетая, само за високото ниво на представяне, съгласно Таблица 12.

Новите, по-тесни, граници на изменение на параметрите на модела са конструирани по следния начин: долната граница се определя да бъде по-малка, но максимално близка до минималната измежду средните стойности на съответния параметър във високото ниво; горната граница се определя да бъде по-голяма, но максимално близка до максималната измежду средните стойности във високото ниво.

В Таблица 14 са представени използваните преди процедурата *широки граници* за всеки едни от параметрите на модела, както и новопредложените *тесни граници* въз основа на ПЦПГ.

Таблица 12 Нива на представяне на СГА_СКМ

СГА_СКМ	Целева функция		Нива на представяне		Усреднено време за сходимост
GGAP = 0.9	<i>min J</i>	0.0221	ВН_ДГ	0.0221	81.6693
			ВН_ГГ	-0.0222	
	<i>avrg J</i>	0.0222	СН_ДГ	0.0222	
			СН_ГГ	0.0222	
	<i>max J</i>	0.0223	НН_ДГ	0.0222	
			НН_ГГ	0.0223	
GGAP = 0.8	<i>min J</i>	0.0221	ВН_ДГ	0.0221	70.3386
			ВН_ГГ	-0.0222	
	<i>avrg J</i>	0.0222	СН_ДГ	0.0222	
			СН_ГГ	0.0222	
	<i>max J</i>	0.0223	НН_ДГ	0.0222	
			НН_ГГ	0.0223	
GGAP = 0.67	<i>min J</i>	0.0221	ВН_ДГ	0.0221	59.1654
			ВН_ГГ	-0.0222	
	<i>avrg J</i>	0.0222	СН_ДГ	0.0222	
			СН_ГГ	0.0222	
	<i>max J</i>	0.0223	НН_ДГ	0.0222	
			НН_ГГ	0.0223	
GGAP = 0.5	<i>min J</i>	0.0222	ВН_ДГ	0.0222	46.9997
			ВН_ГГ	-0.0223	
	<i>avrg J</i>	0.0223	СН_ДГ	0.0223	
			СН_ГГ	-0.0225	
	<i>max J</i>	0.0226	НН_ДГ	0.0225	
			НН_ГГ	0.0226	

Таблица 13 Оценени параметри за високото ниво на представяне на
СГА СКМ

СГА_СКМ		μ_{2S}	μ_{2E}	k_S	k_E	γ_{SX}	γ_{EX}	$k_L^{O_2} a$	γ_{OS}	γ_{OE}
GGAP = 0.9	min	0.94	0.14	0.13	0.80	0.39	1.81	40.42	333.03	35.73
	max	0.99	0.15	0.14	0.80	0.40	2.00	95.53	785.10	96.73
	avrg	0.97	0.14	0.13	0.80	0.39	1.92	63.24	515.78	61.78
GGAP = 0.8	min	0.91	0.12	0.12	0.79	0.40	1.63	52.20	415.38	102.70
	max	1.00	0.14	0.13	0.80	0.41	1.84	126.44	990.90	201.39
	avrg	0.95	0.13	0.13	0.80	0.40	1.78	96.82	769.11	155.37
GGAP = 0.67	min	0.91	0.12	0.11	0.80	0.39	1.67	48.98	379.71	30.46
	max	0.97	0.15	0.14	0.80	0.41	1.84	126.78	987.88	149.45
	avrg	0.95	0.14	0.13	0.80	0.40	1.98	101.39	801.98	92.02
GGAP = 0.5	min	0.91	0.11	0.11	0.79	0.39	1.47	99.59	768.66	102.84
	max	1.00	0.15	0.14	0.80	0.40	2.04	126.78	983.37	261.13
	avrg	0.95	0.14	0.13	0.80	0.40	1.84	108.41	853.07	216.27

Таблица 14 Граници на изменение на параметрите на модела за
СГА СКМ

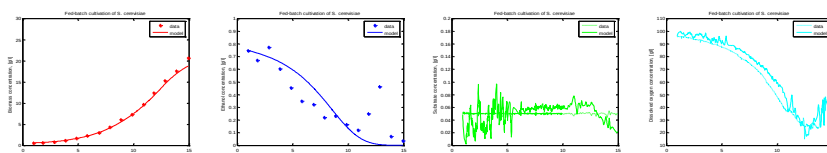
СГА_СКМ		μ_{2S}	μ_{2E}	k_S	k_E	γ_{SX}	γ_{EX}	$k_L^{O_2} a$	γ_{OS}	γ_{OE}
Граници преди ПЦПГ	ДГ	0.9	0.05	0.08	0.5	0.3	1	0.001	0.001	0.001
	ГГ	1	0.15	0.15	0.8	10	10	300	1000	1000
Граници след ПЦПГ	ДГ	0.94	0.13	0.12	0.7	0.38	1.7	60	500	60
	ГГ	0.97	0.15	0.14	0.8	0.42	2	110	900	220

С цел демонстрация на целесъобразността на процедурата за целенасочен параметричен генезис, изследваният СГА_СКМ е изпълнен отново при GGAP = 0.5, с граници на изменение на параметрите, определени след приложение на ПЦПГ. Еднопопулационният генетичен алгоритъм е пуснат неколккратно с цел да бъдат получени надеждни резултати. В Таблица 15 са представени усреднените стойности на целевата функция, времето за сходимост на алгоритъма, както и оценените стойности на параметрите на модела.

Таблица 15 Резултати от параметричната идентификация със СГА_СКМ

Параметър/ GGAP = 0.5	СГА_СКМ	
	Преди приложение на ПЦПГ	След приложение на ПЦПГ
J	0.0222	0.0221
CPU time, s	46.9997	29.25
μ_{2S} , 1/h	0.95	0.95
μ_{2E} , 1/h	0.14	0.14
k_{S1} , g/l	0.13	0.13
k_{E1} , g/l	0.80	0.80
Y_{S1} , g/g	0.40	0.40
Y_{E1} , g/g	1.84	1.87
$k_{L O_2}$, 1/h	108.41	77.21
Y_{O_2} , g/g	853.07	608.53
$Y_{O_2 E}$, g/g	216.27	212.56

Както е показано в Таблица 15, приложението на процедурата за целенасочен параметричен генезис води до намаляване на времето за сходимост на алгоритъма с 38%, без това да повлияе върху точността на модела, което показва ефективността на предложената процедура. На Фиг. 4 са показани резултатите от експерименталните и моделни данни, респективно за концентрациите на биомаса, етанол, субстрат и разтворен кислород.



Фиг. 4 Експериментални и моделни данни за концентрациите на основните биохимични променливи

Приложение на процедура за целенасочен параметричен генезис при модифициран СГА

Ефективността на ПЦПГ е изследвана и за една от модификациите на стандартния еднопопулационен генетичен алгоритъм, а именно модифицирания еднопопулационен ГА, означен като СГА_МКС (*мутация, кръстосване, селекция*).

Целесъобразността на ПЦПГ е тествана посредством неколккратно изпълнение на СГА_МКС, с цел да бъдат получени надеждни резултати. Таблица 16 представя усреднените стойности на целевата функция, времето за сходимост на алгоритъма и параметрите на модела когато СГА_МКС е изпълнен при GGAP = 0.5. Приложението на процедурата за целенасочен параметричен генезис води до намаляване с 29% на времето за сходимост на алгоритъма, при запазване точността на модела.

Таблица 16 Резултати от параметричната идентификация със СГА_МКС

Параметър/ GGAP = 0.5	СГА_МКС	
	Преди приложение на ПЦПГ	След приложение на ПЦПГ
J	0.0224	0.0221
CPU time, s	40.5261	28.6090
μ_{2S} , 1/h	0.92	0.91
μ_{2E} , 1/h	0.12	0.11
k_{S1} , g/l	0.12	0.11
k_{E1} , g/l	0.79	0.80
Y_{SX1} , g/g	0.40	0.42
Y_{EX1} , g/g	1.58	1.51
$k_L^{O_2} a$, 1/h	90.23	69.23
Y_{OS1} , g/g	708.44	550.75
Y_{OE1} , g/g	203.38	194.26

Както се вижда от Таблицы 15 и 16, СГА_МКС е толкова точен колкото е и СГА_СКМ, като едновременно с това неговото време за намиране на решение е далеч по-малко.

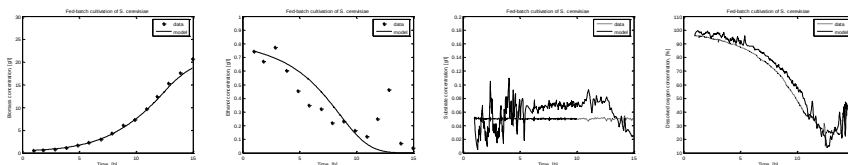
Приложение на процедура за целенасочен параметричен генезис при МПГА

Разработената процедура за целенасочен генезис е приложена и за параметрична идентификация на полупериодична култивация на дрожди *S. cerevisiae*, със стандартен многопопулационен генетичен алгоритъм (МПГА_СКМ). В резултат на приложението на процедурата времето за сходимост на алгоритъма намалява с повече от 12%, като същевременно се наблюдава запазване на точността на намереното решение (Таблица 17).

Таблица 17 Резултати от параметричната идентификация с МПГА_СКМ

Параметър/ GGAP = 0.5	МПГА_СКМ	
	Преди приложение на ПЦПГ	След приложение на ПЦПГ
J	0.0221	0.0220
CPU time, s	98.9596	86.5150
μ_{2S} , 1/h	0.91	0.90
μ_{2E} , 1/h	0.12	0.14
k_{S1} , g/l	0.15	0.15
k_{E1} , g/l	0.80	0.80
Y_{SX1} , g/g	0.41	0.40
Y_{EX1} , g/g	1.62	1.93
$k_L^{O_2} a$, 1/h	96.34	88.73
Y_{OS1} , g/g	768.61	696.56
Y_{OE1} , g/g	809.90	291.42

Фиг. 5 представя резултатите от експерименталните и моделно предсказаните данни след процедурата за целенасочен параметричен генезис с МПГА_СКМ, съответно за биомаса, етанол, субстрат и разтворен кислород.



Фиг. 5 Експериментални и моделни данни за концентрациите на основните биохимични променливи

Резултатите от Таблица 17 и Фиг. 5 показват високата ефективност на предложената процедура за целенасочен параметричен генезис при МПГА_СКМ.

ИЗВОДИ

Постъпкова процедура за целенасочен параметричен генезис за идентификация на параметрите на модела на полупериодична култивация на *S. cerevisiae* е разработена и приложена за два вида еднопопулационни генетични алгоритми. Поради факта, че стандартният СГА_СКМ, е най-добре изучен, той е използван като тестов алгоритъм. Получените многообещаващи резултати са потвърдени чрез приложение на процедурата при модифицирания СГА_МКС с променена последователност на изпълнение на генетичните оператори, а именно *мутация*, *кръстосване* и *селекция*. Прилагането на ПЦПГ води до стесняване на интервалите на изменение на параметрите на модела, като по този начин подобрява работата на ГА. В резултат на това СГА_СКМ става с 38% по-бърз, докато СГА_МКС спестява до 29% от времето за намиране на решение. Посочените резултати показват значително подобряване на бързодействието и на двата алгоритъма, без това да повлияе върху точността на модела. Освен това, ако се сравнят резултатите за СГА_СКМ при GGAP = 0.9 с тези получени със СГА_МКС при GGAP = 0.5, след ПЦПГ, се вижда, че до 65% от времето за сходимост на алгоритъма може да бъде спестено, без това да повлияе върху точността на намереното решение.

Целесъобразността на ПЦПГ е установена и при друг вид ГА, а именно стандартния многопопулационен генетичен алгоритъм. След приложение на ПЦПГ за МПГА_СКМ времето за сходимост на алгоритъма намалява с повече от 12%, като едновременно с това не се наблюдава промяна в точността на намереното решение.

Следователно, основното предимство на предложената процедура за целенасочен параметричен генезис се изразява в това, че прилагането и води до значително подобряване на изпълнението на изследваните алгоритми. Наблюдаваното подобряване се заключава в запазване на постигнатата точност на намереното решение като едновременно с това чувствително се намалява времето за сходимост на изследваните алгоритми.

Удачно е също така да се отбележи, че предложената процедура е универсален инструмент и може да бъде подходящо и успешно приложена и за други стохастични алгоритми за оптимизация, както и за различни обекти, свързани с параметрична идентификация на модел.

Глава 4

Оценяване качеството на представяне на генетичните алгоритми с приложение на интуиционистки размита логика

Ефективността на алгоритъма е изключително важна характеристика, която се определя от сходимостта към глобалния минимум (при многоекстремална целева функция); скоростта на сходимост; броят изчисления на целевата функция; точността на решението (точност на оценката на параметрите на модела); времето за решение на задачата; заеманата компютърна памет, и др. Като представителни критерии за оценка качеството на представяне на даден алгоритъм за целите на параметричната идентификация, могат да бъдат разглеждани, например, физическия смисъл на получените параметри, стойността на целевата функция, бързодействието на алгоритъма, и др. Интуиционистки размитата логика се явява като възможна алтернатива за оценяване качеството на представяне на алгоритмите, като за определяне на степените на вярност и невярност са необходими изпълнения на всеки един от изследваните алгоритми в два различни интервала на изменение на параметрите на модела. По подразбиране единият интервал се идентифицира с т. нар. *широки граници*. Другият интервал, който ще бъде наричан *тесни граници* би могъл да се дефинира по различни начини - въз основа на минималните и максимални стойности на параметрите, след многократни изпълнения на алгоритмите; въз основа на усреднени стойности на параметрите, след прилагане на процедурата за целенасочен параметричен генезис, както и въз основа на други идеи в зависимост от спецификата на конкретния проблем.

Процедура за оценяване качеството на представяне на генетични алгоритми чрез приложение на интуиционистки размита логика

В интуиционистки размитата логика [31, 32], ако p е променлива, то нейната вярностна стойност се представя чрез наредената двойка:

$$V(p) = \langle M(p), N(p) \rangle \quad (8)$$

където $M(p), N(p), M(p) + N(p) \in [0, 1]$, а $M(p)$ и $N(p)$ са съответно степените на вярност и на невярност на p . Стойностите на тези степени могат да бъдат получени с помощта на различни формули в зависимост от разглеждания проблем. В [10, 11] е дефинирана зависимост от вида $\leq$ между интуиционистки размитите двойки $\langle a, b \rangle$ и $\langle c, d \rangle$, чрез

$$\langle a, b \rangle \leq \langle c, d \rangle \text{ тогава и само тогава когато } a \leq c \text{ и } b \geq d.$$

За целите на настоящото изследване степените на вярност/невярност могат да бъдат получени, например, по следния начин:

$$M(p) = \frac{m}{u}, N(p) = 1 - \frac{n}{u} \quad (9)$$

където m е долната от *тесните* граници на изменение на даден параметър; u – горната от *широките* граници на изменение на даден параметър; n – горната от *тесните* граници на изменение на съответния оценен параметър.

Ако се натрупва база от данни, състояща се от елементи от вида $\langle p, M(p), N(p) \rangle$, могат да бъдат получени различни нови оценки на променливите, като например *оптимистична*, *средна* и *песимистична*. В случай на два записа в базата данни, новите оценки могат да бъдат получени съгласно следните формули:

- оптимистична

$$V_{opt}(p) = \langle \max(M_1(p), M_2(p)), \min(N_1(p), N_2(p)) \rangle, \quad (10)$$

- средна

$$V_{aver}(p) = \langle ((M_1(p) + M_2(p))/2, (N_1(p) + N_2(p))/2) \rangle, \quad (11)$$

- песимистична

$$V_{pes}(p) = \langle \min(M_1(p), M_2(p)), \max(N_1(p), N_2(p)) \rangle. \quad (12)$$

Следователно, за всяко p

$$V_{pes}(p) \leq V_{aver}(p) \leq V_{opt}(p).$$

В случай на три записа в базата данни, новите оценки – съответно *строго оптимистична*, *оптимистична*, *средна*, *песимистична* и *строго песимистична* могат да бъдат получени съгласно следните формули:

- строго оптимистична

$$V_{strong_opt}(p) = \langle M_1(p) + M_2(p) + M_3(p) - M_1(p)M_2(p) - M_1(p)M_3(p) - M_2(p)M_3(p) + \\ + M_1(p)M_2(p)M_3(p), N_1(p)N_2(p)N_3(p) \rangle, \quad (13)$$

- оптимистична

$$V_{opt}(p) = \langle \max(M_1(p), M_2(p), M_3(p)), \min(N_1(p), N_2(p), N_3(p)) \rangle, \quad (14)$$

- средна

$$V_{aver}(p) = \langle (M_1(p) + M_2(p) + M_3(p))/3, (N_1(p) + N_2(p) + N_3(p))/3 \rangle, \quad (15)$$

- песимистична

$$V_{pes}(p) = \langle \min(M_1(p), M_2(p), M_3(p)), \max(N_1(p), N_2(p), N_3(p)) \rangle, \quad (16)$$

- строго песимистична

$$V_{strong_pes}(p) = \langle M_1(p)M_2(p)M_3(p), N_1(p) + N_2(p) + N_3(p) - N_1(p)N_2(p) - N_1(p)N_3(p) - \\ - N_2(p)N_3(p) + N_1(p)N_2(p)N_3(p) \rangle \quad (17)$$

Следователно, за всяко p

$$V_{strong_pes}(p) \leq V_{pes}(p) \leq V_{aver}(p) \leq V_{opt}(p) \leq V_{strong_opt}(p).$$

Като алтернатива за оценка качеството на представяне на даден алгоритъм по отношение на стойността на целевата функция и бързодействието на алгоритъма е разработена процедура, чрез приложение на ИРЛ. Процедурата се състои от шест стъпки:

- Стъпка 1.* Извършват се определен брой изпълнения на всеки един от алгоритмите, обект на конкретното изследване, за двата интервала на изменение на параметрите на модела – *широки* и *тесни*.
- Стъпка 2.* Определят се усреднени стойности на целевата функция, времето за сходимост на алгоритъма и оценяваните параметри.
- Стъпка 3.* Определят се степените на вярност и невярност за всеки от алгоритмите, обект на конкретното изследване, съгласно (9).
- Стъпка 4.* Определят се границите за *оптимистична*, *средна* и *песимистична* оценка на параметрите на модела, в случая на сравнение на два алгоритъма, съгласно (10)-(12) и *строго оптимистичната*, *оптимистичната*, *средната*, *песимистичната* и *строго песимистичната* оценка, в случая на сравнение на три алгоритъма, съгласно (13)-(17).
- Стъпка 5.* Поставяне на оценките от *Стъпка 4* за всеки параметър, за всеки интервал на изменение и за всеки от разглежданите алгоритми.

Стъпка 6. Оценяване качеството на представяне на всеки един от алгоритмите на базата на получените в Стъпка 5 оценки.

Приложение на процедурата за оценяване качеството на представяне на генетични алгоритми

Оценяване качеството на представяне на различни видове СГА

Извършено е сравнение по отношение качеството на представяне на три вида СГА, а именно стандартния, означен като СГА_СКМ и на две негови модификации СГА_КСМ и СГА_МСК. Преди да бъде пристъпено към приложението на процедурата за оценяване качеството на представяне на трите алгоритъма, за всеки от тях са извършени по тридесет изпълнения в т. нар. *широки граници*, на базата, на които да бъдат определени новите, „по-тесни“ граници (Стъпка 1 от ПОКП). В конкретния случай *тесните граници* на изменение са определени въз основа на минималните и максимални стойности на оценяваните параметри. Новите, „по-тесни“ граници са конструирани по следния начин: долната граница се определя така, че да бъде по-малка, но максимално близка до минималната стойност на съответния параметър; горната граница се определя така, че да бъде по-голяма, но максимално близка до максималната стойност на съответния параметър. Таблица 18 представя двата интервала на изменение на параметрите на модела с долни (ДГ) и горни граници (ГГ) както и изчислените степени на вярност/невярност (Стъпка 3 от ПОКП) за трите алгоритъма, получени въз основа на (9).

Таблица 18. Граници на изменение и степени на вярност/невярност за различните видове СГА

			μ_{2S}	μ_{2E}	k_S	k_E	γ_{SX}	γ_{EX}	$k_L^{O_2} a$	γ_{OS}	γ_{OE}
СГА_СКМ	Широки граници	ДГ	0.90	0.05	0.08	0.50	0.30	1.00	0.001	0.001	0.001
		ГГ	1.00	0.15	0.15	0.80	10.00	10.00	300.00	1000.00	1000.00
	Тесни граници	ДГ	0.91	0.07	0.09	0.60	0.30	1.00	50.00	300.00	100.00
		ГГ	1.00	0.15	0.15	0.80	0.50	3.00	130.00	900.00	250.00
	Степен на вярност на p	$M_1(p)$	0.91	0.47	0.60	0.75	0.03	0.10	0.17	0.30	0.10
	Степен на невярност на p	$N_1(p)$	0.00	0.00	0.00	0.00	0.95	0.70	0.57	0.10	0.75
СГА_КСМ	Широки граници	ДГ	0.90	0.05	0.08	0.50	0.30	1.00	0.001	0.001	0.001
		ГГ	1.00	0.15	0.15	0.80	10.00	10.00	300.00	1000.00	1000.00
	Тесни граници	ДГ	0.91	0.10	0.09	0.60	0.30	1.20	120.00	800.00	20.00
		ГГ	0.95	0.14	0.15	0.80	0.50	2.50	200.00	1000.00	130.00
	Степен на вярност на p	$M_2(p)$	0.91	0.67	0.60	0.75	0.03	0.12	0.40	0.80	0.02
	Степен на невярност на p	$N_2(p)$	0.05	0.07	0.00	0.00	0.95	0.75	0.33	0.00	0.87
СГА_МСК	Широки граници	ДГ	0.90	0.05	0.08	0.50	0.30	1.00	0.001	0.001	0.001
		ГГ	1.00	0.15	0.15	0.80	10.00	10.00	300.00	1000.00	1000.00
	Тесни граници	ДГ	0.90	0.08	0.09	0.70	0.30	1.00	50.00	300.00	100.00
		ГГ	0.94	0.13	0.14	0.80	0.50	2.00	130.00	900.00	250.00
	Степен на вярност на p	$M_3(p)$	0.90	0.53	0.60	0.88	0.03	0.10	0.17	0.30	0.10
	Степен на невярност на p	$N_3(p)$	0.06	0.13	0.07	0.00	0.95	0.80	0.57	0.10	0.75

Таблица 19 представя границите за *строго оптимистичната*, *оптимистичната*, *средната*, *песимистичната* и *строго песимистичната* оценка за стойностите на параметрите на модела (Стъпка 4 от ПОКП) за трите вида СГА, получени въз основа на (9) и формули (13)-(17).

Таблица 19. Оценки за различните видове СГА

		μ_{2S}		μ_{2E}		k_S		k_E		Y_{SX}		Y_{EX}		$k_L^{O_2} a$		Y_{OS}		Y_{OE}	
		ДГ	ГГ	ДГ	ГГ	ДГ	ГГ	ДГ	ГГ	ДГ	ГГ	ДГ	ГГ	ДГ	ГГ	ДГ	ГГ	ДГ	ГГ
Различни видове СГА	$V_{strong\ opt}$	1.00	1.00	0.14	0.15	0.14	0.15	0.79	0.80	0.87	1.43	2.87	5.80	175.00	267.89	902.00	1000.00	206.20	510.63
	V_{opt}	0.91	1.00	0.10	0.15	0.09	0.15	0.70	0.80	0.30	0.50	1.20	3.00	120.00	200.00	800.00	1000.00	100.00	250.00
	V_{aver}	0.91	0.96	0.08	0.14	0.09	0.15	0.63	0.80	0.30	0.50	1.07	2.50	73.33	153.33	466.67	933.33	73.33	210.00
	V_{pes}	0.90	0.94	0.07	0.13	0.09	0.14	0.60	0.80	0.30	0.50	1.00	2.00	50.00	130.00	300.00	900.00	20.00	130.00
	$V_{strong\ pes}$	0.75	0.89	0.02	0.12	0.03	0.14	0.39	0.80	0.00	0.00	0.01	0.15	3.33	37.56	72.00	810.00	0.20	8.12

Изследваните три вида СГА са приложени отново за целите на параметрична идентификация на полупериодичен процес за култивация на дрождите *S. cerevisiae* като са използвани границите от Таблица 18. За да се гарантира получаване на достоверни резултати трита вида СГА са изпълнени по тридесет пъти. Таблица 20 представя усреднените стойности на целевата функция, времето за сходимост на алгоритъма, както и оценените стойности на параметрите на модела с трита вида СГА, съответно в *широки* и *тесни граници*.

Таблица 20. Резултати от параметрична идентификация с различните видове СГА

	СГА_СКМ		СГА_КСМ		СГА_МСК	
	<i>широки граници</i>	<i>тесни граници</i>	<i>широки граници</i>	<i>тесни граници</i>	<i>широки граници</i>	<i>тесни граници</i>
J	0.0222	0.0221	0.0228	0.0223	0.0223	0.0222
$CPU\ time, s$	47.00	29.25	36.22	32.00	37.69	30.83
$\mu_{2S}, 1/h$	0.95	0.95	0.91	0.93	0.91	0.91
$\mu_{2E}, 1/h$	0.14	0.14	0.12	0.14	0.11	0.13
$k_S, g/l$	0.13	0.13	0.09	0.13	0.11	0.12
$k_E, g/l$	0.8	0.8	0.70	0.75	0.80	0.80
$Y_{SX}, g/g$	0.4	0.4	0.42	0.38	0.44	0.40
$Y_{EX}, g/g$	1.84	1.87	2.01	1.90	1.53	1.69
$k_L^{O_2} a, 1/h$	108.41	77.21	125.39	121.06	121.27	81.07
$Y_{OS}, g/g$	853.07	608.53	875.91	955.45	874.40	694.74
$Y_{OE}, g/g$	216.27	212.56	98.84	111.23	120.58	127.29

Трябва да бъде отбелязано, че и при трита вида СГА изпълнението на алгоритмите в *тесни граници* води до очаквано намаляване на времето за сходимост като се запазва постигнатата точност на модела. Изпълняването на СГА в *тесни граници* намалява времето за достигане до решението с повече от 1.6 пъти при СГА_СКМ и съответно с 1.13 и 1.22 пъти при приложението на СГА_КСМ и СГА_МСК.

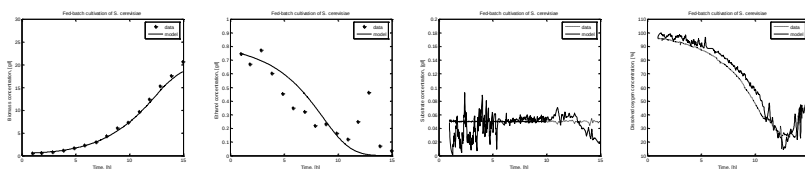
В Таблица 21 са представени оценките, съгласно Таблица 19, за всеки един от параметрите, получени с всеки един от разглежданите СГА, изпълнени съответно в *широки* и *тесни граници*. Както се вижда от Таблица 21, няма нито една получена строго песимистична и песимистична прогноза. В два от случаите има три строго оптимистични прогнози, следвани съответно от пет и четири оптимистични прогнози – това са случаите на СГА_СКМ в *широки граници* и на СГА_СКМ в *тесни граници*. Много близки до тези резултати се получават при изпълнение на СГА_КСМ в *тесни граници* – две строго

оптимистични и седем оптимистични, но без нито една средна прогноза. В тези три случая, отличени като най-надеждни, стойността на обектната функция е близка до нейната най-ниска стойност, което означава, че изследваните алгоритми достигат висока степен на точност на намереното решение. Ако трите отличени алгоритъма се сравнят по време за бързодействие, то най-бърз е СГА_СКМ в *тесни граници*. Той е 1.61 пъти по-бърз от СГА_СКМ в *широки граници* и дори е малко по-бърз – 1.09 пъти от СГА_КСМ в *тесни граници*. Дори и с една оптимистична оценка по-малко от СГА_СКМ в *широки граници* СГА_СКМ в *тесни граници* е най-бърз и е с достигната най-висока степен на точност на модела. Ето защо този алгоритъм може да бъде отличен като най-надежден, тъй като целта е да бъдат получени резултати с висока степен на точност за най-кратко изчислително време.

Таблица 21. Оценки за резултатите от параметрична идентификация с различните видове СГА

	СГА_СКМ		СГА_КСМ		СГА_МСК	
	<i>широки граници</i>	<i>тесни граници</i>	<i>широки граници</i>	<i>тесни граници</i>	<i>широки граници</i>	<i>тесни граници</i>
строго оптимистична	3	3	0	2	0	1
оптимистична	5	4	8	7	8	6
средна	1	2	1	0	1	2
песимистична	0	0	0	0	0	0
строго песимистична	0	0	0	0	0	0

Фиг. 6 представя получените резултати за експерименталните и моделни данни за концентрациите, съответно на биомаса, етанол, субстрат и разтворен кислород когато СГА_СКМ е приложен в *тесни граници* за целите на параметрична идентификация на дрождите *S. cerevisiae*.



Фигура 6. Експериментални и моделни данни за концентрациите на основните биохимични променливи при приложение на СГА_СКМ в *тесни граници*

Оценяване качеството на представяне на различни видове МПГА

Посредством предложената ПОКП е извършено сравнение на качеството на представяне на два вида МПГА – стандартния, означен като МПГА_СКМ и една негова модификация, означена като МПГА_МСК. В конкретното изследване *тесните граници* на изменение на параметрите на модела са получени след прилагане на процедурата за целенасочен параметричен генезис.

Таблица 22. Резултати от параметрична идентификация с различните видове МПГА

	МПГА_СКМ		МПГА_МКС	
	<i>широки граници</i>	<i>тесни граници</i>	<i>широки граници</i>	<i>тесни граници</i>
J	0.0221	0.022	0.0221	0.0221
CPU time, s	98.9596	86.515	270.95	245.42
$\mu_{2S}, 1/h$	0.91	0.90	0.93	0.90
$\mu_{2E}, 1/h$	0.12	0.14	0.13	0.13
$k_S, g/l$	0.15	0.15	0.15	0.15
$k_E, g/l$	0.80	0.80	0.80	0.80
$Y_{SX}, g/g$	0.41	0.40	0.40	0.41
$Y_{EX}, g/g$	1.62	1.93	1.81	1.77
$k_L^{O_2}, 1/h$	96.34	88.73	62.59	84.41
$Y_{OS}, g/g$	768.61	696.56	500.75	673.20
$Y_{OE}, g/g$	809.90	291.42	454.40	462.89

Трябва да се отбележи, че и при двата разглеждани МПГА изпълнението им в *тесни граници* води до очаквано намаляване на времето за сходимост. Изпълнението на МПГА в *тесни граници* намалява времето за намиране на решение с 1.14 и 1.10 пъти когато са приложени съответно МПГА_СКМ и МПГА_МКС. В допълнение към гореизложеното може да се каже, че получените в *тесни граници* резултати попадат в най-високото постигнато до момента ниво на точност на полученото решение, което показва добрата ефективност на всички разглеждани тук МПГА по отношение на тяхната сходимост и точност на намиране на решението.

Таблица 23 съдържа оценките за всеки един от параметрите, получени с всеки един от разглежданите МПГА в *широки* и *тесни граници*.

Таблица 23. Оценки за резултатите от параметричната идентификация с различните видове МПГА

	МПГА_СКМ		МПГА_МКС	
	<i>широки граници</i>	<i>тесни граници</i>	<i>широки граници</i>	<i>тесни граници</i>
оптимистична	9	9	7	9
средна	0	0	0	0
песимистична	0	0	2	0

Както се вижда от Таблица 23, в три от четирите случая има получени само оптимистични прогнози, което демонстрира високата надеждност и на двата МПГА. Единственият случай с прогноза, различна от оптимистичната е случаят с МПГА_МКС в *широки граници*. Следователно в този случай не съществува друг начин за сравняване на двата алгоритъма освен по времето за сходимост. Тъй като, МПГА_СКМ е най-бърз, в *тесни граници* на изпълнение, той може да бъде считан за „лидер“ измежду отличените алгоритми. МПГА_СКМ в *тесни граници* на изпълнение е 1.14 пъти по-бърз от същия изпълнен в *широки граници*. Но ако бъдат сравнени МПГА_СКМ, в *тесни граници* на изпълнение, с МПГА_МКС то той се явява 3.13 пъти по-бърз от МПГА_МКС в *широки граници* и 2.84 пъти по-бърз от МПГА_МКС в *тесни граници*. Този факт води до логичното заключение, че МПГА_СКМ в *тесни граници* може да бъде отличен като най-надеждния алгоритъм, в случай че се

цели получаване на резултати с висока степен на точност за минимално изчислително време.

Оценяване качеството на представяне на стандартен СГА спрямо стандартен МПГА

ПОКП е приложена и за сравнение на качеството на представяне на стандартните СГА_СКМ и МПГА_СКМ. Въз основа на натрупването на данни са определени новите *тесни граници* на изменение на параметрите за всеки алгоритъм, съгласно минималната и максимална стойност за всеки един от параметрите.

Таблица 24 представя усреднените стойности на целевата функция, времето за сходимост на алгоритъма, оценените стойности на параметрите на модела с двата вида ГА и съответстващите им интуиционистки размити оценки.

Таблица 24. Оценка на резултатите от приложението на СГА_СКМ и МПГА_СКМ

	СГА_СКМ		МПГА_СКМ	
	<i>широки граници</i>	<i>тесни граници</i>	<i>широки граници</i>	<i>тесни граници</i>
J	0.0222	0.0221	0.0221	0.022
CPU time, s	47.00	29.25	98.96	86.52
μ_{2S} , 1/h	0.95 opt	0.95 opt	0.91 opt	0.9 pes
μ_{2E} , 1/h	0.14 opt	0.14 opt	0.12 opt	0.14 opt
k_S , g/l	0.13 aver	0.13 aver	0.15 opt	0.15 opt
k_E , g/l	0.8 opt	0.8 opt	0.8 opt	0.8 opt
Y_{SX} , g/g	0.4 opt	0.4 opt	0.41 opt	0.4 opt
Y_{EX} , g/g	1.84 opt	1.87 opt	1.62 opt	1.93 opt
$k_L^{O_2} a$, 1/h	108.41 opt	77.21 aver	96.34 opt	88.73 opt
Y_{OS} , g/g	853.07 opt	608.53 aver	768.61 opt	696.56 opt
Y_{OE} , g/g	216.27 aver	212.56 aver	809.9 opt	291.42 opt

Въз основа на Таблица 24 могат да бъдат направени следните изводи: Съгласно интуиционистки размитите оценки, МПГА_СКМ в *широки граници* е най-надежден, тъй като всички стойности на параметрите са с оптимистична оценка. Най-близко до МПГА_СКМ в *широки граници* е МПГА_СКМ в *тесни граници* с осем оптимистични и една песимистична оценка. Но и при двата интервала на изменение на параметрите МПГА_СКМ е чувствително по-бавен от СГА_СКМ: СГА_СКМ в *широки граници* е 2.11 пъти по-бърз от МПГА_СКМ в *широки граници* и съответно, СГА_СКМ в *тесни граници* е дори 2.96 пъти по-бърз от МПГА_СКМ в *тесни граници*. В този случай изборът на надежден алгоритъм остава въпрос на предпочитание на изследователя. Като разумен компромис може да бъде отличен СГА_СКМ в *широки граници*, тъй като е втори по бързодействие, но пак е много по-надежден от най-бързия СГА_СКМ в *тесни граници*.

Оценяване качеството на представяне на стандартен СГА и МПГА при различни стойности на GGAP

В настоящата дисертация GGAP е отличен като най-чувствителният параметър на ГА по отношение времето за сходимост на алгоритъма. Затова качеството на представяне на стандартен СГА и стандартен МПГА, означени съответно като СГА_СКМ и МПГА_СКМ, е изследвано за три различни стойности на GGAP, а именно $GGAP = 0.9$, $GGAP = 0.5$ и $GGAP = 0.1$.

Качеството на представяне на СГА_СКМ е изследвано преди и след приложение на ПЦПГ за всяка от трите стойности на GGAP. Таблица 25 представя усреднените стойности на целевата функция, времето за сходимост на алгоритъма, както и оценените стойности на параметрите на модела чрез приложение на СГА_СКМ при различни стойности на GGAP, съответно в *широки* и *тесни граници*.

Таблица 25. Резултати от параметрична идентификация със СГА_СКМ за различни стойности на GGAP

	СГА_СКМ за GGAP = 0.9		СГА_СКМ за GGAP = 0.5		СГА_СКМ за GGAP = 0.1	
	<i>широки граници</i>	<i>тесни граници</i>	<i>широки граници</i>	<i>тесни граници</i>	<i>широки граници</i>	<i>тесни граници</i>
J	0.0222	0.0221	0.0223	0.0221	0.0225	0.0222
CPU time, s	81.67	43.82	46.99	30.48	25.98	18.45
μ_{2S} , 1/h	0.94	0.95	0.94	0.95	0.95	0.95
μ_{2E} , 1/h	0.12	0.15	0.12	0.14	0.11	0.12
k_S , g/l	0.12	0.13	0.13	0.13	0.13	0.12
k_E , g/l	0.80	0.80	0.75	0.80	0.75	0.79
Y_{SX} , g/g	0.41	0.40	0.40	0.40	0.41	0.40
Y_{EX} , g/g	1.44	2.00	1.58	1.92	1.48	1.67
$k_i^{O_2}$, 1/h	70.28	70.32	109.27	108.31	79.31	94.03
Y_{OS} , g/g	544.37	553.71	858.75	864.15	627.78	729.57
Y_{OE} , g/g	102.31	59.01	156.92	155.49	187.09	192.43

Таблица 26 представя оценките, за всеки един от параметрите на модела, получени след приложение на СГА_СКМ при трите различни стойности на GGAP в *широки* и *тесни граници*, съответно преди и след ПЦПГ.

Таблица 26. Оценки за резултатите от параметрична идентификация със СГА_СКМ при различни стойности на GGAP

	СГА_СКМ за GGAP = 0.9		СГА_СКМ за GGAP = 0.5		СГА_СКМ за GGAP = 0.1	
	<i>широки граници</i>	<i>тесни граници</i>	<i>широки граници</i>	<i>тесни граници</i>	<i>широки граници</i>	<i>тесни граници</i>
строго оптимистична	1	2	1	2	1	1
оптимистична	4	4	6	7	4	4
средна	4	3	2	0	4	4
песимистична	0	0	0	0	0	0
строго песимистична	0	0	0	0	0	0

Както се вижда от Таблица 26, няма нито една *строго песимистична* и *песимистична* оценка, което още веднъж потвърждава ефективността на

използвания СГА_СКМ. Анализът показва, че в два от общо шестте случая има по две *строго оптимистични* прогнози – това са СГА_СКМ в *тесни граници* при GGAP = 0.9 и СГА_СКМ в *тесни граници* при GGAP = 0.5. В другите четири случая представянето се оценява с по една *строго оптимистична* оценка. Броят на *оптимистичните* и *средните* оценки е сравним, с лек превес на *оптимистичните* оценки в случая на СГА_СКМ в *широки граници* при GGAP = 0.5. Очевидно е, че има три абсолютно еднакви представяния на алгоритъма: при GGAP = 0.9 в *широки граници* и при GGAP = 0.1, съответно в *широки* и *тесни граници*. Във всички случаи, при които са използвани получените *тесни граници* за параметрите на модела, стойността на целевата функция намалява и е много близка да най-ниската такава, което пакзва, че е достигната висока степен на точност на модела. Измежду отличените две представяния на СГА_СКМ с по две *строго оптимистични* оценки, СГА_СКМ в *тесни граници* при GGAP = 0.5 е оценен със седем *оптимистични* оценки, докато СГА_СКМ в *тесни граници* при GGAP = 0.9 е оценен съответно с четири *оптимистични* и три *средни* оценки. Следователно, въз основа на интуиционистки размитите оценки за параметрите на модела, СГА_СКМ при GGAP = 0.5 в *тесни граници* е отличен като най-надежден алгоритъм, в случай че се цели висока точност за най-малко изчислително време. Но ако целта е бързодействие на алгоритъма, най-бързият измежду разгледаните е СГА_СКМ при GGAP = 0.1 в *тесни граници*.

Качеството на представяне на другия алгоритъм - МПГА_СКМ е изследвано преди и след приложение на ПЦПГ. В Таблица 27 са представени усреднените стойности на целевата функция, времето за сходимост на алгоритъма, както и оценените стойности на параметрите на модела с МПГА_СКМ, съответно в *широки* и *тесни граници* за трите различни стойности на GGAP.

Таблица 27. Резултати от параметрична идентификация с МПГА_СКМ за различни стойности на GGAP

	МПГА_СКМ за GGAP = 0.9		МПГА_СКМ за GGAP = 0.5		МПГА_СКМ за GGAP = 0.1	
	<i>широки границы</i>	<i>тесни границы</i>	<i>широки границы</i>	<i>тесни границы</i>	<i>широки границы</i>	<i>тесни границы</i>
J	0.0221	0.221	0.0221	0.0221	0.0222	0.0221
CPU time, s	159.67	148.91	98.96	90.51	34.16	32.13
$\mu_{2S}, 1/h$	0.91	0.90	0.91	0.91	0.94	0.91
$\mu_{2E}, 1/h$	0.12	0.13	0.13	0.13	0.11	0.13
$k_S, g/l$	0.15	0.15	0.15	0.15	0.14	0.15
$k_E, g/l$	0.80	0.80	0.80	0.80	0.79	0.80
$Y_{SX}, g/g$	0.41	0.41	0.41	0.41	0.42	0.41
$Y_{EX}, g/g$	1.64	1.77	1.69	1.76	1.45	1.70
$k_E^{P_2} \cdot a, 1/h$	84.24	93.74	94.40	89.37	95.57	88.86
$Y_{OS}, g/g$	669.65	742.89	742.47	708.43	743.15	708.15
$Y_{OE}, g/g$	334.89	404.11	516.92	468.28	458.77	475.75

Приложението на МПГА_СКМ в *тесни граници*, получени след прилагане на ПЦПГ, намалява от 6 до 10% времето за сходимост на алгоритъма, като запазва точността на модела. Както се вижда от Таблица 28 няма нито една *строго песимистична* и *песимистична* оценка. В четири от случаите има по три *строго оптимистични* оценки, а в три от тях следващите

шест оценки са *оптимистични* – това са случаите при $GGAP = 0.5$ преди и след ПЦПГ и $GGAP = 0.1$ след ПЦПГ. В тези три случая, отличени по гореописания начин като най-надеждни, стойността на целевата функция съвпада с най-ниската получавана до момента стойност на целевата функция, което отговаря на най-високата постигната степен на точност. Но ако трита случая бъдат сравнени по отношението на времето за намиране на решение, МПГА_СКМ с $GGAP = 0.1$ в *тесни граници* е около 3 пъти по-бърз от МПГА_СКМ с $GGAP = 0.5$. Следователно, въз основа на интуиционистки размитите оценки за параметрите на модела, МПГА_СКМ с $GGAP = 0.1$ след ПЦПГ, е отличен като най-надежден алгоритъм, в случай че целта е получаване на висока точност за най-малко изчислително време.

Таблица 28. Оценки за резултатите от параметрична идентификация с МПГА_СКМ при различни стойности на $GGAP$

	МПГА_СКМ за $GGAP = 0.9$		МПГА_СКМ за $GGAP = 0.5$		МПГА_СКМ за $GGAP = 0.1$	
	<i>широки граници</i>	<i>тесни граници</i>	<i>широки граници</i>	<i>тесни граници</i>	<i>широки граници</i>	<i>тесни граници</i>
<i>строго оптимистична</i>	2	3	3	3	1	3
<i>оптимистична</i>	7	5	6	6	7	6
<i>средна</i>	0	1	0	0	1	0
<i>песимистична</i>	0	0	0	0	0	0
<i>строго песимистична</i>	0	0	0	0	0	0

ИЗВОДИ

В Глава 4 интуиционистки размитата логика е използвана за оценка качеството на представяне на различни видове генетични алгоритми. ИРП е приложена за оценка качеството на представяне на различни видове СГА и МПГА, както и за оценяване на влиянието на $GGAP$, доказан като най-чувствителния параметър на ГА по отношение времето за намиране на решение. Всички разгледани ГА са приложени за целите на параметрична идентификация на модел на култивиране на дрождите *S. cerevisiae*.

Най-напред, ПОКП е приложена за оценка качеството на представяне на различни видове СГА, а именно на стандартния СГА_СКМ и на две негови модификации - СГА_КСМ и СГА_МСК. СГА_СКМ в *тесни граници* на изменение на параметрите е отличен като най-бърз – 1.61 пъти по-бърз в сравнение със СГА_СКМ в *широки граници* и 1.09 пъти по-бърз в сравнение със СГА_КСМ в *тесни граници*. Дори и с една оптимистична прогноза по-малко от СГА_СКМ в *широки граници*, СГА_СКМ в *тесни граници* е най-бърз и с най-висока постигната точност на модела, следователно може да се отличи като най-надежден алгоритъм, ако се цели получаване на резултати с висока степен на точност за най-малко изчислително време.

След това ПОКП е приложена за оценка качеството на представяне на различни видове МПГА, а именно стандартния МПГА_СКМ и модифицирания МПГА_МСК. Заслужава да се отбележи, че в три от разгледаните четири случая резултатите от идентификацията с МПГА получават само оптимистични оценки, което потвърждава надеждността на използваните видове МПГА за параметрична идентификация на модел на

ферментационен процес. Като най-бърз алгоритъм от всички разгледани МПГА може да бъде отличен МПГА_СКМ в *тесни граници*.

ПОКП е приложена и за оценка точността на намерените решения със стандартен СГА и стандартен МПГА. На пръв поглед МПГА_СКМ в *широки граници* е най-надежден, защото всички намерени с него стойности на параметрите на модела получават оптимистични оценки. Но и в двата случая, съответно в *тесни* и *широки граници* на изменение на моделните параметри МПГА_СКМ са значително по-бавни, отколкото СГА_СКМ. Като приемлив компромис може да се приеме СГА_СКМ в *широки граници* на изменение, въпреки че е втори по бързодействие, тъй като този алгоритъм е по-надежден от най-бързият СГА_СКМ в *тесни граници*.

Също така ПОКП е приложена за оценка качеството на представяне на СГА и МПГА при различни стойности на GGAP, който е отличен като най-чувствителния по отношение на бързодействието параметър на ГА. В два от случаите СГА_СКМ са оценени с по две *строго оптимистични* прогнози - СГА_СКМ в *тесни граници* при GGAP = 0.5 и СГА_СКМ, отново в *тесни граници*, но при GGAP = 0.9. Останалите прогнози за оценените параметри, са съответно 7 *оптимистични* при GGAP = 0.5, както и 4 *оптимистични* и 3 *средни* при GGAP = 0.9. Въз основа на това, СГА_СКМ при GGAP = 0.5 в *тесни граници* на изпълнение е отличен като най-надежден алгоритъм, в случай, че се цели да бъдат получени резултати с висока точност за по-малко изчислително време. Но ако целта е да бъде намалено времето за изчисление на алгоритъма, най-бърз сред изследваните се оказва СГА_СКМ в *тесни граници* при GGAP = 0.1.

Влиянието на GGAP е изследвано и при МПГА. След приложение на ИРЛ за получаване на интуиционистки размити оценки за параметрите на модела и последващото определяне на *строго оптимистичните*, *оптимистичните*, *средните*, *песимистичните* и *строго песимистичните* прогнози за резултатите, получени с МПГА_СКМ, при трите различни стойности на GGAP са сравнени и МПГА_СКМ при GGAP = 0.5 в *широки* и *тесни граници* както и МПГА_СКМ при GGAP = 0.1 след ПЦПГ са отличени идентично. Измежду излъчените три лидера, МПГА_СКМ с GGAP = 0.1 след ПЦПГ е около три пъти по-бърз от МПГА_СКМ при GGAP = 0.5 като едновременно с това запазва достигнатата висока точност на модела.

В заключение може да се каже, че представените тук „кръстосани оценки“ на различни видове СГА и МПГА, на стандартен СГА спрямо стандартен МПГА, както и за влиянието на най-чувствителния параметър на ГА по отношение на тяхното бързодействие, доказват приложимостта на интуиционистки размитите оценки като помощен инструмент при оценяване качеството на представяне на генетичните алгоритми.

Следва да се отбележи, че предложената процедура е универсална и може да бъде успешно приложена както при параметрична идентификация на различни обекти, така и за други алгоритми за оптимизация.

ГЛАВА 5

Заключение

Въз основа на представените резултати, приносите в дисертационния труд могат да бъдат обобщени както следва:

1. Разработени са пет нови модификации на СГА, различаващи се по реда на изпълнение на основните генетични оператори. Три от новопредложените модификации, а именно СГА_МКС, СГА_МСК и СГА_КСМ, водят до по-добра сходимост на алгоритъма при параметрична идентификация на култивация на дрожди, без това да влияе на точността на намереното решение. [A2, A5]
2. Разработени са шест нови модификации на МПГА, различаващи се по реда на изпълнение на основните генетични оператори. Три от новопредложените модификации, а именно МПГА_КСМ, МПГА_МСК и МПГА_СМК, водят до по-добра сходимост на алгоритъма при параметрична идентификация на култивация на дрожди, без това да влияе на точността на намереното решение. [A1, A5]
3. Предложена е оптимална настройка на параметрите на СГА и МПГА след изследване на влиянието на параметрите върху сходимостта на алгоритмите и точността на намереното решение. И при двата вида ГА най-чувствителен по отношение на бързодействието на алгоритмите се оказва параметърът степен на прехвърляне. С близо 40% може да бъде подобро времето за сходимост на СГА и МПГА, без това да се отрази върху точността на намереното решение. [A1, A2, A5]
4. Разработена е процедура за целенасочен генезис на параметрите на модела. Приложението на процедурата при параметрична идентификация на култивация на дрожди води до значително намаляване на времето за сходимост на алгоритмите, съответно с малко над 10% при МПГА_СМК, близо 30% при СГА_МКС и близо 40% при СГА_СМК, без това да се отрази върху точността на модела. Предложената процедура е универсална и може да бъде успешно приложена както при параметрична идентификация на различни обекти, така и за други алгоритми за оптимизация. [A4]
5. Разработена е процедура с приложение на ИРП за оценка качеството на представяне на ГА. Процедурата е успешно приложена при параметрична идентификация на култивация на дрожди за оценка на различни видове СГА, различни видове МПГА, стандартни СГА и МПГА и ГА с различни стойности на степента на прехвърляне. Предложената процедура е универсална и може да бъде успешно приложена както при параметрична идентификация на различни обекти, така и за други алгоритми за оптимизация. [A3]

Публикации и апробация на резултатите в дисертационния труд

- A1. **Angelova M.**, T. Pencheva, Improvement of Multi-population Genetic Algorithms Convergence Time, *In Monte Carlo Methods and Applications*, Edited by Sabelfeld, Karl K./ Dimov, Ivan, Chapter 1, ISBN: 9783110293586, De Gruyter, Berlin, Germany, 2013, 1-10.
- A2. **Angelova M.**, T. Pencheva, Algorithms Improving Convergence Time in Parameter Identification of Fed-batch Cultivation, *Comptes rendus de l'Académie bulgare des Sciences*, 2012, 65(3), 299-306.
- A3. **Ангелова М.**, Т. Пенчева, Процедура за сравнение качеството на генетични алгоритми чрез приложение на интуиционистки размита логика, *X Национална младежка научно-практическа сесия*, София,

2012, 244-249.

- A4. **Ангелова М.**, Т. Пенчева, Целенасочен параметричен генезис с многопопулационен генетичен алгоритъм, *X Национална младежка научно-практическа сесия*, София, 2012, 250-254.
- A5. **Angelova M.**, Т. Pencheva, Tuning Genetic Algorithm Parameters to Improve Convergence Time, *International Journal of Chemical Engineering*, vol. 2011, Article ID 646917, 7 pages, 2011. doi:10.1155/2011/646917.

ЛИТЕРАТУРА

1. Илкова Т., Моделиране и оптимизация на клас ферментационни процеси. Докторска дисертация, София, 2007.
2. Косев К., О. Роева, Модифициран многопопулационен генетичен алгоритъм с подобрена сходимост. 8ма Национална младежка научно-практическа сесия, София, България, 2010, 7-12.
3. Миланов П., Н. Янев, И. Мирчев, И. Тренчев, Теория на алгоритмите и оптимизация. ЮЗУ Неофит Рилски, 2010.
4. Основни елементи на съвременната биотехнология, 2009. (<http://www.referati.org/osnovni-elementi-nasyvremennatabiotehnologiq/35620/ref>).
5. Пенчева Т., Моделиране на клас биотехнологични процеси като обекти с разпределени параметри. Докторска дисертация, София, 2003.
6. Роева О., Моделиране на функционални състояния при култивиране на *E. coli* с приложение на генетични алгоритми. Докторска дисертация, София, 2007.
7. Цонков Ст. и кол., Основи на автоматичното управление на биотехнологични процеси, София, 1994.
8. Цонков Ст. и кол., Управление на биотехнологичните процеси, София, Техника, 1992.
9. Цонков Ст., Й. Костов, М. Петров, Т. Пенчева, Пл. Златева, Кр. Лякова, Т. Илкова, Я. Христозов, Б. Илиев, О. Роева, Ю. Христова, Биопроцесни системи: Моделиране, управление и оптимизация, Изток-Запад, София, 2004.
10. Atanassov K., Intuitionistic fuzzy sets. Heidelberg, Springer, 1999.
11. Atanassov K., On intuitionistic fuzzy sets theory. Berlin, Springer, 2012.
12. Chande S., M. Sinha, Genetic Algorithm: A Versatile Optimization Tool. BVICAM'S International Journal of Information Technology, New Delhi, 2008, 7-13.
13. Chipperfield A. J., P. Fleming, H. Pohlheim, C. M. Fonseca, Genetic Algorithm Toolbox for Use with MATLAB, User's guide, version 1.2. Dept. of Automatic Control and System Engineering, University of Sheffield, UK, 1994.
14. Chipperfield A. J., P. J. Fleming, The Matlab Genetic Algorithm Toolbox, 1995.
15. Goldberg D., Genetic Algorithms in Search, Optimization and Machine Learning, Wiley, Massachusetts, 1989.
16. Gupta D., S. Ghafir, An overview of methods maintaining diversity in genetic algorithms. International Journal of Emerging Technology and Advanced Engineering, 2(5), 2012, 56-60.
17. Hristozov I., Т. Pencheva, St. Tzonkov, B. Hitzmann, Functional State Modelling Approach for Batch Cultivation of *Saccharomyces cerevisiae*, Chemical and Biochemical Engineering Quarterly, 19(1), 2005, 69-74.
18. <http://winnie-the-pooh.hit.bg/mathlogic.html>

19. Kenneth J., K. Dhinakar, Cybernetic Model of the Growth Dynamics of *Saccharomyces Cerevisiae* in Batch and Continuous Cultures, *Jornal of Biotechnology*, 71, 1999, 105-131.
20. Leao C., F. Soares, Heuristic Sensitivity Analysis for Baker's Yeast Model Parameters, *Investigacao Operacional*, 24, 2004, 247-263.
21. Neumaier A., Global Optimization. <http://solon.cma.univie.ac.at/~neum/glopt.html>, 2003.
22. Obittko M., Genetic Algorithm. (2005), <http://cs.felk.cvut.cz/~xobittko/ga/main.html>
23. Papagianni M., Y. Boonpooh, M. Mattey, B. Kristiansen, Substrate Inhibition Kinetics of *Saccharomyces cerevisiae* in Fed-batch Cultures Operated at Constant Glucose and Maltose Concentration Levels, *Journal of Industrial Microbiology and Biotechnology*, 34, 2007, 301–309.
24. Pencheva T., I. Hristozov, D. Huell, B. Hitzmann, St. Tzonkov, Modelling of Functional States during *Saccharomyces cerevisiae* Fed-batch Cultivation, *Bioautomation*, 2, 2005, 8-16.
25. Pencheva T., O. Roeva, I. Hristozov, Functional State Approach to Fermentation Processes Modelling, Prof. Marin Drinov Academic Publishing House, Sofia, 2006.
26. Pham H., G. Larsson, S. Enfors, Modelling of Aerobic Growth of *Saccharomyces Cerevisiae* in a pH-auxostat, *Bioprocess Engineering*, 20, 1999, 537-544.
27. Pohlheim H., Genetic and Evolutionary Algorithms: Principles, Methods and Algorithms, http://www.systemtechnik.tu-ilmenau.de/~pohlheim/GA_Toolbox/algindex.html
28. Roeva O., A Modified Genetic Algorithm for a Parameter Identification of Fermentation Processes, *Biotechnology and Biotechnological Equipment*, 20(1), 2006, 202-209.
29. Roeva O., Sensitivity Analysis of *E. coli* Fed-batch Cultivation Local Models, *Mathematica Balkanica*, 4, 2011, 395-413.
30. Rowe J., I. East, Direct replacement: A Genetic Algorithm without Mutation which Avoids Deception, *Lecture notes in Artificial Intelligence*, Springer, 956, 1995, 41-48.
31. Rychtera M., L. Paulova, J. Nahlik, K. Melzoch, J.Votuba, Control Strategy of Fed-Batch Cultivations of Yeasts, *Chemical Papers*, 50(4), 1996, 238-244.
32. Vicente A., M. Dluhy, E. Ferreira, J. Teixeira, Modelling Diffusion-Reaction Phenomena in Yeast Flocs of *Saccharomyces Cerevisiae*, *Bioprocess Engineering*, 18, 1998, 335-342.
33. Vilums S., O. Grigs, Application of Functional State Modelling Approach for Yeast *Saccharomyces Cerevisiae* Batch Fermentation State Estimation, *International Conference on Applied Information and Communication Technologies (AICT2012)*, 26.-27. April, 2012, Jelgava, Latvia <http://aict.itf.llu.lv300> .
34. Zhang X.-C., A. Visala, A. Halme, P. Linco, Functional State Modelling Approach for Bioprocesses: Local Models for Aerobic Yeast Growth Processes, *Journal of Process Control*, 4(3), 1994, 127-134.