DISCRETE-TIME SYSTEM IDENTIFICATION USING GENETIC ALGORITHM WITH SINGLE PARENT-BASED MATING

FARAH AYIESYA BINTI ZAINUDDIN

DOCTOR OF PHILOSOPHY

2024

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

Faculty of Mechanical Technology and Engineering

DISCRETE-TIME SYSTEM IDENTIFICATION USING GENETIC ALGORITHM WITH SINGLE PARENT-BASED MATING TECHNIQUE

Farah Ayiesya binti Zainuddin

Doctor of Philosophy

2024

DISCRETE-TIME SYSTEM IDENTIFICATION USING GENETIC ALGORITHM WITH SINGLE PARENT-BASED MATING TECHNIQUE

FARAH AYIESYA ZAINUDDIN

A thesis submitted
in fulfilment of the requirements for the degree of Doctor of Philosophy

Faculty of Mechanical Technology and Engineering

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2024

DECLARATION

I declare that this thesis entitled "Discrete-Time System Identification using Genetic Algorithm with Single Parent-Based Mating Technique" is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature : ……………………………………

Name : Farah Ayiesya Binti Zainuddin

Date : 18 October 2024

APPROVAL

I hereby declare that I have read this thesis and in my opinion this thesis is sufficient in terms of scope and quality for the award of Doctor of Philosophy.

Signature	:	………………………………………
Supervisor Name	:	Assoc. Prof. Ir. Dr. Md Fahmi Bin Abd Samad @ Mahmood ………………………………………
Date	:	18 October 2024 ………………………………………

DEDICATION

In the name of Allah, the Most Gracious and the Most Merciful

I dedicate this work to:

My parents,

Zainuddin bin Mohd Noh and Maznah binti Md Yusoff

My husband and my daughter,

Muhammad Amir Asyraf bin Abd Kadir and Nur Rania Tihani binti Muhammad Amir Asyraf,

My siblings,

Mentor who always give support and encouragement,

Assoc. Prof. Ir. Dr. Md Fahmi Bin Abd Samad @ Mahmood

ABSTRACT

System identifiction (SI) is a methodology for developing mathematical models of dynamic systems using measurements of input and output signals. This research focuses on improving Genetic Algorithms (GA) for SI, specifically addressing inefficiencies in common crossover operators that limit search space exploration and lead to premature convergence. The study aims to enhance the performance of GA by introducing a novel Single Parent Mating (SPM) technique. The objectives are to enhance the performance of crossover operator for GA, simulate SI using GA with the SPM technique, and compare the performance of the modified GA to traditional GA in terms of prediction accuracy, convergence to global optimum, and model parsimony. The methodology encompasses data acquisition, GA program development, SPM technique implementation, and simulation using MATLAB. The study simulated single-input-single-output (SISO) models: ARX and NARX. Performance indicators included the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Parameter Magnitude-based Information Criterion 2 (PMIC2). The findings showed that incorporating SPM with single-point crossover resulted in lower objective function (OF) values and improved error indices (EI) compared to traditional GA methods. This enhanced the GA's ability to avoid premature convergence and maintained a diverse solution set, leading to more optimal model selection. The study's results were validated using real-world data from industrial systems, including a hair dryer, an air compression system, and a flexible robot arm. In these cases, the SPM technique consistently outperformed traditional GA, demonstrating improved model fit and predictive accuracy. Rigorous validation tests, including autocorrelation and cross-correlation functions, confirmed the reliability and robustness of these models. These practical applications underscore the versatility and effectiveness of the SPM technique in enhancing GA for SI, proving its utility across different fields such as engineering, finance, and healthcare. The successful validation in real-world systems marks a significant milestone, showing that the SPM technique can significantly improve model optimization in diverse and practical contexts . This research makes several important contributions to the field of GA and SI. The introduction of the SPM technique represents a significant enhancement in the performance of GA, particularly in terms of improving genetic diversity and search optimization . The findings provide a robust framework for applying GAs to complex modeling problems, emphasizing the importance of effective crossover strategies and genetic diversity . The SPM technique offers researchers and practitioners a powerful tool for achieving faster convergence, better optimization, and more accurate models. This study not only advances the theoretical understanding of GA but also provides practical methodologies that can be applied to real-world problems, making it a valuable contribution to both academic research and practical applications.

PENGENALPASTIAN SISTEM DISKRET-MASA MENGGUNAKAN ALGORTIMA GENETIK DENGAN TEKNIK PENGAWANAN BERASASKAN INDUK TUNGGAL

ABSTRAK

Pengenalpastian sistem (SI) adalah metodologi untuk membangunkan model matematik sistem dinamik menggunakan pengukuran isyarat masukan dan keluaran. Penyelidikan ini memberi tumpuan kepada memperbaiki Algoritma Genetik (GA) untuk SI, dengan menangani ketidakcekapan dalam operator penyilang biasa yang mengehadkan penerokaan ruang carian dan membawa kepada penumpuan awal. Kajian ini bertujuan untuk meningkatkan prestasi GA dengan memperkenalkan teknik Pengawanan Induk Tunggal (SPM) yang baru. Objektif kajian ini adalah untuk menigkatkan prestasi operator penyilang yang lebih baik untuk GA, mensimulasikan SI menggunakan GA dengan teknik SPM, dan membandingkan prestasi GA yang diubah suai dengan GA tradisional dari segi ketepatan ramalan, penumpuan kepada optimum global, dan kesederhanaan model. Metodologi kajian merangkumi pemerolehan data, pembangunan program GA, pelaksanaan teknik SPM, dan simulasi menggunakan MATLAB. Kajian ini mensimulasikan model masukan-tunggal-keluaran-tunggal (SISO): ARX dan NARX. Penunjuk prestasi termasuk Kriteria Maklumat Akaike (AIC), Kriteria Maklumat Bayesian (BIC), dan Kriteria Maklumat berasaskan Magnitud Parameter 2 (PMIC2). Hasil kajian menunjukkan bahawa penggabungan SPM dengan penyilang titik tunggal menghasilkan nilai fungsi objektif (OF) yang lebih rendah dan indeks ralat (EI) yang lebih baik berbanding dengan kaedah GA tradisional. Ini meningkatkan keupayaan GA untuk mengelakkan penumpuan awal dan mengekalkan set penyelesaian yang pelbagai, yang membawa kepada pemilihan model yang lebih optimum. Keputusan kajian ini disahkan menggunakan data dunia sebenar dari sistem industri, termasuk pengering rambut, sistem pemampatan udara, dan lengan robot fleksibel. Dalam kes ini, teknik SPM secara konsisten mengatasi GA tradisional, menunjukkan kesesuaian model dan ketepatan ramalan yang lebih baik. Ujian pengesahan yang ketat, termasuk fungsi autokorelasi dan silang-korelasi, mengesahkan kebolehpercayaan dan ketahanan model ini. Aplikasi praktikal ini menekankan kepelbagaian dan keberkesanan teknik SPM dalam meningkatkan GA untuk SI, membuktikan kegunaannya dalam pelbagai bidang seperti kejuruteraan, kewangan, dan penjagaan kesihatan. Pengesahan berjaya dalam sistem dunia sebenar menandakan pencapaian penting, menunjukkan bahawa teknik SPM boleh meningkatkan pengoptimuman model dalam pelbagai konteks praktikal. Penyelidikan ini memberikan beberapa sumbangan penting kepada bidang GA dan SI. Pengenalan teknik SPM mewakili peningkatan yang ketara dalam prestasi GA, terutamanya dari segi peningkatan kepelbagaian genetik dan pengoptimuman carian. Penemuan ini menyediakan rangka kerja yang kukuh untuk menerapkan GA kepada masalah pemodelan yang kompleks, menekankan kepentingan strategi penyilangan yang berkesan dan kepelbagaian genetik. Teknik SPM menawarkan alat yang kuat kepada penyelidik dan pengamal untuk mencapai penumpuan yang lebih cepat, pengoptimuman yang lebih baik, dan model yang lebih tepat. Kajian ini bukan sahaja memajukan pemahaman teori tentang GA tetapi juga menyediakan metodologi praktikal yang boleh diterapkan kepada masalah

dunia sebenar, menjadikannya sumbangan berharga kepada penyelidikan akademik dan aplikasi praktikal.

ACKNOWLEDGEMENTS

First and foremost, I thank Allah s.w.t, the Almighty God for His mercy and grace in enabling me to complete this thesis work. I would like to take this opportunity to express my sincere acknowledgement to my supervisor Assoc. Prof. Dr. Md Fahmi Bin Abd Samad @ Mahmood from the Faculty of Mechanical Technology and Engineering, Universiti Teknikal Malaysia Melaka (UTeM) for their essential supervision, support and encouragement towards the completion of this thesis.

Special thanks to the Kementerian Pengajian Tinggi (KPT) under the funding of Fundamental Research Grant Scheme (FRGS), Faculty of Mechanical Technology and Engineering and Centre for Advanced Computing Technology, Universiti Teknikal Malaysia Melaka (UTeM) (FRGS/1/2018/TK03/UTEM/02/13) for financial sponsorship, facilities and knowledge during this research.

Finally, I would like to make an affectionate acknowledgement to my family, all my colleagues and friends who had helped and supported me during my research studies directly or indirectly.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACF	-	Autocorrelation function
AI	-	Artificial intelligence
AIC	-	Akaike information criterion
ANN	-	Artificial neural network
ARX	-	AutoRegressive with eXogenous input
ARMAX	-	AutoRegressive Moving Average with eXogenous input
ASS	-	Assembly sequence scheduling
BIC	-	Bayesian infromation criterion
BJ	-	Box-Jenkins
CDA	-	Compressed dry air
CCF	-	Cross-correlation function
DC	-	Direct current (constant value in mathematical model)
EA	-	Evolutionary algorithms
EC	-	Evolutionary computation
EI	-	Error index
ES	-	Evolutionary strategies
ESS	-	Explained sum of squares
FF	-	Fitness function
GA	-	Genetic algorithm
HLCMGA	-	Crossover with human-like constrained mating
HW	-	Hammerstein-Wiener
IBAX	-	Inversed Bi-segmented average crossover
LSS	-	Linear state-space
ML	-	Machine learning
MSE	-	Mean square error
MWh	-	Megawatt-hour
NARX	-	Nonlinear AutoRegressive with eXogenous input
NARMAX	-	Nonlinear AutoRegressive Moving Average with eXogenous inputs

NLSS	-	Nonlinear state-space
OE	-	Output error
OF	-	Objective function
OLS	-	Ordinary least squares
PF	-	Penalty function
PMIC2	-	Parameter magnitude based information criterion 2
PSO	-	Particle swarm optimization
RSS	-	Residual sum of square
SCADA	-	Supervisory control and data acquisition
SI	-	System identification
SISO	-	Single-input-single-output
SPM	-	Single parent mating
TF	-	Transfer function
TSP	-	Traveling salesman problem
TSS	-	Total sum of squares

LIST OF SYMBOLS

a_i, b_i and c_i	-	Parameter values of regressor i
C	-	Constant
d	-	Time delay
$e(t)$	-	Noise value at time t
$E(P)$	-	Entropy (population)
$F_*^1[.]$	-	Nonlinear function
$fitness_i$	-	Fitness value for chromosome i
k (in k-point)	-	Number of crossover point
k (in k-step ahead)	-	Smallest lag order of predicted output in a difference equation model
L	-	Number of regressors
l	-	Degree of nonlinearity
lchrom	-	Length of bit string chromosome
M	-	Number of independent variables
n	-	Number of individuals in the population
n_y	-	Output lag order
n_u,	-	Input lag order
n_k	-	System delay
N	-	Total number of data
N (in N-point)	-	Number of crossover point
P	-	Partition according to fitness value
popsize	-	Population size
p_c	-	Crossover probability or rate
p_i	-	Selection probability of individual i

p_m	-	Mutation probability or rate
p_k	-	partition k
R^2	-	Squared multiple correlation coefficient
t	-	Time; Target output
$u(t)$	-	Input signal
x	-	Rate of electricity; Data sets of input
$X \triangleq (x_1, x_2 .. x_n)$	-	Sample vector
$\chi 2$	-	Chi-squared test
$y(t)$	-	Output value at time t
$\hat{y}(t)$	-	Predicted output value at time t
$Y \triangleq (y_1, y_2 .. y_n)$	-	Corresponding label of vector X on n dimensional space
Z^N	-	Function of parameter vector
$\varepsilon(t)$	-	Residual or prediction error at time t
$\emptyset$	-	Regresor vector
$\emptyset_{\varepsilon\varepsilon}$	-	Autocorrelation function
$\emptyset_{u\varepsilon}$	-	Cross-correlation function
$\emptyset_{u^2\varepsilon}$	-	Correlation square of the input and residuals
$\emptyset_{u^2\varepsilon^2}$	-	Correlation square of the input and square of the residuals
$\emptyset_{\varepsilon\varepsilon u}$	-	Correlation residuals and product of the residuals and input
θ	-	True parameter vector

LIST OF APPENDICES

LIST OF PUBLICATIONS

Zainuddin, F.A. and Samad, M.F.A., 2020. A Review of Crossover Methods and Problem Representation of Genetic Algorithm in Recent Engineering Applications. *International Journal of Advanced Science and Technology*, 29(6s), pp. 759-769.

Zainuddin, F.A. and Samad, M.F.A., 2021. Comparison of Crossover in Genetic Algorithm for Discrete-Time System Identification. *International Review of Mechanical Engineering (IREME)*, 15(2), pp. 59-66.

Zainuddin, F.A. and Samad, M.F.A., 2021. A Mating Technique for Various Crossover in Genetic Algorithm for Optimum System Identification. *International Review of Mechanical Engineering (IREME)*, 15(11), pp. 574-581.

Samad, M.F.A. and Zainuddin, F.A., 2023. Single Parent Mating in Genetic Algorithm for Real Robotic System Identification. *International Journal of Artificial Intelligence (IAES)*, 12(1), pp. 201-208.

Zainuddin, F.A. and Samad, M.F.A., 2020. Crossover of Genetic Algorithm for Linear and Nonlinear System Identification. *Proceeding of Innovative Research and Industrial Dialogue (IRID'20)*, pp. 145-146.

Zainuddin, F.A. and Samad, M.F.A., 2021. Optimum System Identification Incorporating Mating Technique in Genetic Algorithm. *Proceedings of Malaysian Technical Universities Conference on Engineering and Technology (MUCET) 2021*, pp. 162-163.

Zainuddin, F.A., Samad, M.F.A., Jamaluddin, H. and Azad, A.K.M., 2024. A Novel Single – Parent Mating Technique in Genetic Algorithm for Discrete – Time System Identification. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, 42(2), pp. 49-57.

Samad, M.F.A and Zainuddin, F.A., 2024. Mating in Genetic Algorithm with Application in System Identification. In: *Khosravy, M., Gupta, N. and Witkowski O. (Eds.) Frontiers in Genetic Algorithm Theory and Application*, pp. 35-43, Springer Tracts in Nature–Inspired Computing, Springer.