NEO X

10th International Workshop on
Numerical and Evolutionary Optimization
November 8-10 2022

General Chair:

Marcela Quiroz Artificial Intelligence Research Institute, UV

Contents

Xalapa

Foreword

Welcome

Welcome to NEO 2022 (NEO X), the 10th International Workshop on Numerical and Evolutionary Optimization. In this edition, NEO X occurs from November 08 to 10, 2022. As in the last year, the workshop will be held as online only event.

The goal of the Numerical and Evolutionary Optimization (NEO) workshop series is to bring together people from these and related fields to discuss, compare and merge their complementary perspectives. NEO encourages the development of fast and reliable hybrid methods, that maximize the strengths and minimize the weaknesses of each underlying paradigm; while also applying to a broader class of problems. Moreover, NEO fosters the understanding and adequate treatment of real-world problems, particularly in emerging fields that affect us all such as healthcare, smart cities, big data, among many others.

In this 2022 edition, there will be more than 80 participants coming from all over Mexico. The NEO X will have more than 50 technical presentations addressing different optimization subjects and dealing with a variety of challenging applications. It is worth to notice that over the years, the NEO community has grown not only in number but also concerning gender diversity. This year we arrange also, as now it is common since 2018, a special session called Women at NEO (W-NEO) as an effort to make more visible the work in optimization performed by our female pairs. Finally, the NEO X will hold RED again, the *Research Experience Day*, where undergraduate students can attend conferences, tutorials, and discussion panels focused on awakening their interest in the wide variety of research career opportunities available to them.

We hope you enjoy your participation at the NEO 2022, thank you for your valuable assistance.

Sincerely,

Marcela Quiroz, Universidad Veracruzana

NEO 2022 General Chair

Acknowledgments

We want to thank all participants that helped to make the NEO 2022 (NEO X) such a great success. In particular, we would like to thank Keynote Speakers:
Michael Emmerich, Leiden University, The Netherlands;
Lee Spector, Amherst College, USA;
Laura Cruz Reyes, ITCM of TecNM, Mexico; and
Una-May O'Reilly, MIT, USA.

as well as the Speakers of the Research Experience Day 2022:
Carlos Ignacio Hernández Castellanos, IIMAS-UNAM, México;
Guadalupe Carmona Arroyo, Universidad Veracruzana, México; and
Enrique Naredo García, Universidad de Limerick, Irlanda.

Further, we thank to the institutions
CONACYT,
Cinvestav-IPN,
Universidad Veracruzana,
Tecnológico Nacional de México/IT Tijuana, in Tijuana, and
Tecnológico Nacional de México/IT de Cd. Madero, in Cd. Madero;

to the academic colleagues
Dr. Daniel E. Hernández Morales (TECNM/IT Tijuana)
Oliver Cuate (ESFM-IPN),
Guillermo Morales (Cinvestav-IPN),
Brisbane Ovilla (Cinvestav-IPN),
Claudia Gpe. Gómes Santillán (TECNM/IT de Cd. Madero),
María Lucila Morales Rodríguez (TECNM/IT de Cd. Madero),
Cuautehmoc Mancilla López (Cinvestav-IPN);

and to the staff members
Marco Antonio Ortega Flores (IPN),
Katia Ocampo (Cinvestav-IPN),
Santiago Dominguez (Cinvestav-IPN),
Erika Rios (Cinvestav-IPN),
Jose-Luis Flores (Cinvestav-IPN),
Adriana Martinez (Kreaprom SA de CV and Impakt 45 SA de CV), and
Everett Zhu (MDPI).

NEO 2022 Organizing Committee

Partners

Cinvestav

Universidad Veracruzana

IPN

"POR MI PATRIA Y POR MI BIEN"

TIJUANA, BAJA CALIFORNIA

CONACYT
Consejo Nacional de Ciencia y Tecnología

MDPI

TREE LAB

10th International Workshop on
Numerical and Evolutionary Optimization

NEOX

November 08-10, 2022

Online Only Event

Schedule NEO 2022

Each session ends with a **Session Wrap Up**. Times are 20 minutes per talk, as well as for the wrap ups. All times are in accordance with CDMX time zone.

Day 1, November 08, 2022
Research Experience Day

9:15 – 9:30 Welcome to the RED

9:30 – 10:30 Panel de Aniversario de NEO, "Jóvenes investigadores comparten su experiencia en NEO".
MIA Guadalupe Carmona Arroyo, Dr. Carlos Ignacio Hernández Castellanos, Dr. Enrique Naredo García

10:30 – 11:00 Conferencia: Un vistazo a optimización multi-objetivo y algoritmos evolutivos
Dr. Carlos Ignacio Hernández Castellanos, IIMAS-UNAM, México

11:00 – 11:30 Conferencia: Optimización Inteligente para la Inteligencia Artificial
MIA Guadalupe Carmona Arroyo, Universidad Veracruzana, México

11:30 – 12:00 Conferencia: Diseño Automático de Circuitos Digitales
Dr. Enrique Naredo García, Universidad de Limerick, Irlanda

12:00 – 12:10 Coffee break

12:10 – 13:10 Oportunidades de posgrados
CINVESTAV-IPN, UV, TecNM/ITT, TecNM/ITCM

13:10 – 13:25 Closing

16:00 – 18:00 **Women at NEO**

Day 2, November 09, 2022

09:00 – 09:20 Opening

09:20 – 10:30 Keynote I: **Michael Emmerich**
From Darwin to Newton: Evolutionary and Numerical Methods for Indicator-based Pareto Front Approximation 15

10:30 – 10:40 Coffee break

10:40 – 12:20 Sessions I and II (2 in parallel, up to 4 talks each)

12:20 – 12:30 Coffee break

12:30 – 14:10 Sessions III and IV (2 in parallel, up to 4 talks each)

14:10 – 15:00 Lunch break

15:00 – 16:10 Keynote II: **Lee Spector**
Honor all the things: A lexicase selection manifesto 17
16:20 – 16:30 Coffee break
16:30 – 17:45 Sessions V and VI (2 in parallel, 3 talks each)
17:45 – 17:55 Coffee break
17:55 – 19:10 Sessions VII and VIII (2 in parallel, 3 talks each)

Session I: 10:40 – 12:20, Room 1, **Decision Making & Intelligent Agents 1**
Chair: Claudia Gpe. Gómez Santillán
(ID 0507, in page 25) **Manuel Vargas-Martinez, Nelson Rangel Valdez and Eduardo Fernández** Performance Analysis of Multi-Objective Simulated Annealing Based on Decomposition
(ID 3803, in page 27) **Pablo Hernández-Vicente, Claudia Gpe. Gómez Santillán, Nelson Rangel Valdez and Fausto Balderas-Jaramillo** Multi-objective optimization through an interactive evolutionary framework with uncertainty
(ID 2321, in page 29) **Lorena Rosas-Solórzano, Claudia Gpe. Gómez Santillán, Laura Cruz Reyes and Nelson Rangel Valdez** Coevolutionary algorithm applying outranking method with uncertainty for multi-objective optimization
(ID 8988, in page 31) **José Alfredo Brambila Hernández, Eduardo Villegas-Huerta, Héctor Fraire, Miguel Ángel García-Morales and Armando Becerra de Angel** Hybrid Harmonic Search Algorithm with Improved Opposition-Based Learning (HHS-IOBL) in Benchmark Functions

Session II: 10:40 – 12:20, Room 2, **COVID & Healthcare 1**
Chair: Nelson Rangel Valdez
(ID 0649, in page 34) **Santiago Sinisterra-Sierra, Salvador Godoy-Calderón and Miriam Pescador-Rojas** COVID-19 data analysis with a Multi-objective Evolutionary Algorithm for Causal Association Rules Mining
(ID 1837, in page 35) **Christian Lizbeth Noguez-Moreno, Edwin Montes-Orozco, Ángel David Tellez-Macías, Gilberto Sinuhé Torres-Cockrell, Roman Anselmo Mora-Gutierrez and Miguel Ángel Fernández Romero** Relationship between the development of COVID-19 and poverty in the municipalities of Mexico
(ID 6488, in page 37) **Roberto Herrera-Charles, Teodoro Alvarez-Sanchez and Jesús A. Alvarez-Cedillo** Binaural and multi tone generator to characterize nonlinear systems for healthcare and device manufacturing applications
(ID 7935, in page 38) **Oswaldo Sánchez Andrade, Roman A. Mora-Gutiérrez, Edwin Montes-Orozco, Sergio G. de los Cobos Silva, Eric A. Rincón-García, Pedro Lara-Velázquez, Miguel A. Gutiérrez Andrade and Jorge Juárez Gómez** A comparison of the behavior of the traditional education system and the remote teaching emerging project implemented by Autonomous Metropolitan University into the emergence of COVID-19

Session III: 12:30 – 14:10, Room 1, **Discrete Optimization 1**
Chair: Marcela Quiroz
(ID 1354, in page 40) **José Fernando Padrón-Tristán, Laura Cruz Reyes, Rafael Espin-Andrade and Carlos Eric Llorente-Peralta** A genetic algorithm for predicate

discovery with compensating fuzzy logic

(ID 2158, in page 42) **Jordan Michelt Aran Perez, Laura Cruz Reyes, Bernabé Dorronsoro, Héctor Fraire, Nelson Rangel Valdez, Claudia Gpe. Gómez Santillán and Marcela Quiroz Castellanos** A cooperative coevolutionary genetic approach to solve packing problems

(ID 2189, in page 44) **Jessica González-San-Martín, Laura Cruz Reyes, Héctor Fraire and Bernabé Dorronsoro** Study of Methods for One-dimensional Bin Packing Problem

(ID 3416, in page 46) **Stephanie Amador Larrea, Marcela Quiroz Castellanos and Guillermo-de-Jesús Hoyos Rivera** A new high performance crossover operator for the Grouping Genetic Algorithm with Controlled Gene Transsmision

Session IV: 12:30 – 14:10, Room 2, **Genetic Programming & Machine Learning 1**

Chair: Daniel Hernández

(ID 0340, in page 47) **Erick Estrada Patiño, Juan Frausto-Solis and Guadalupe Castilla-Valdez** Hurricanes ocurrences and temperature forecasting by machine learning methods

(ID 1356, in page 48) **Luis Muñoz Delgado and Adrian Rodriguez Aguiñaga** M3GP data refining under different scopes

(ID 1952, in page 51) **José-Clemente Hernández-Hernández, Gustavo-Adolfo Vargas-Hakim, David Herrera-Sánchez and Efrén Mezura-Montes** Feature Selection with Single Objective Particle Swarm Optimization in Gene Expression Data

(ID 3500, in page 52) **Jose M. Muñoz, Leonardo Trujillo and Daniel E. Hernández** Benchmarking GSGP-CUDA in SRBench

Session V: 16:20 – 17:45, Room 1, **Optimization in Industry 1**

Chair: Oliver Cuate

(ID 0624, in page 53) **Yeudiel Lara Moreno and Carlos Ignacio Hernández Castellanos** A Hierarchical Approach to a Tri-objective Portfolio Optimization Problem Considering an ESG Index

(ID 0753, in page 54) **Julieta Ramiro Estrada, Sebastián Sandoval Montiel, Oliver Cuate and Lourdes Uribe** A multiobjective approach for personalized inversion

(ID 5925, in page 55) **Ricardo Pérez-Rodríguez and Sergio Frausto-Hernández** A radial hybrid estimation of distribution algorithm for the truck and trailer routing problem

Session VI: 16:20 – 17:45, Room 2, **NEO 1**

Chair: Oliver Schütze

(ID 4902, in page 56) **Angel Eduardo Rodríguez-Fernández** Solving Discrete Problems with Continuous Optimization

(ID 1150, in page 57) **Raymundo Juárez del Toro, Vadim Azhmyakov, Carmen Borrego Salcido and Alejandro S. Fonseca Zendejas** A Generalized Linear Model estimation for agricultural variables by an extension of the so-called Recursive Least Square Method

(ID 2276, in page 58) **Oliver Cuate and Oliver Schütze** A Continuation Method for the Treatment of Degenerated Multi-objective Optimization Problems

Session VII: 17:55 – 19:10, Room 1, **Genetic Programming & Machine Learning 2**
Chair: Leonardo Trujillo
(ID 3632, in page 59) **Maritza Talhia Bernabé-Morales, Erasmo Cadenas-Calderón, Rafael Campos-Amezcua and Itzagueri Garcia-Rodriguez** Combination of WRF numerical model and ARIMA model to improve wind speed forecasts
(ID 3724, in page 60) **Jose Luis Purata-Aldaz, Juan Frausto-Solis, Juan Javier González-Barbosa and Guadalupe Castilla-Valdez** Enhanced combinatorial forecasting for Mexican Stock Exchange

Session VIII: 17:55 – 19:10, Room 2, **NEO 2**
Chair: Oliver Schütze
(ID 3071, in page 64) **Luis Ángel Meza-Zárate, Roman Anselmo Mora-Gutiérrez, Edwin Montes-Orozco, Juan Villegas-Cortez, Jorge Mercado-Mondragon, Jorge Juárez-Gómez and Bibiana Obregón Quintana** A bibliometric analysis of texts on homophobia, transphobia, HIV discrimination, and gender violence from 2012 to 2022
(ID 4027, in page 66) **Victor Manuel Sanchez Sanchez, Carlos Gershenson Garcia and Carlos Ignacio Hernández Castellanos** On the Effect of Temporal Heterogeneity on Selection Pressure of Evolutionary Algorithms
(ID 8118, in page 67) **Carlos Hernández and Oliver Schütze** A Bounded Archiver for Hausdorff Approximations of the Pareto Front for Multi-Objective Evolutionary Algorithms

Day 3, November 10, 2022

09:00 – 10:10 Keynote III: **Laura Cruz-Reyes**
Where is research going in multi-objective evolutionary optimization incorporating user preferences? 19
10:10 – 10:20 Coffee break
10:20 – 12:00 Sessions IX and X (2 in parallel, 4 talks each)
12:00 – 12:10 Coffee break
12:10 – 13:50 Sessions XI and XII (2 in parallel, 4 talks each)
13:50 – 15:00 Lunch break
15:00 – 16:10 Keynote IV: **Una-May O'Reilly**
Modeling Adversarial Dynamics 21
16:10 – 16:20 Coffee break
16:20 – 17:45 Sessions XIII and XIV (2 in parallel, 3 talks each)
17:45 – 18:00 Closing

Session IX: 10:20 – 12:00, Room 1, **Discrete Optimization 2**
Chair: Marcela Quiroz
(ID 3779, in page 68) **Sara Mandujano, Adriana Lara and Juan Carlos Ku-Cauich** On the Study of Some Design Aspects for Evolutionary Algorithms for the Generation of Boolean Functions
(ID 4799, in page 69) **Guadalupe Carmona-Arroyo, Marcela Quiroz Castellanos and Efrén Mezura-Montes** An experimental study of the Grouping Genetic Algorithm

Invited Speakers

Michael Emmerich

From Darwin to Newton: Evolutionary and Numerical Methods for Indicator-based Pareto Front Approximation.

Leiden University, The Netherlands

Abstract

Abstract: The concept of indicator-based optimization originated in the field of population-based metaheuristics for multi-objective optimization. The key idea is to measure the quality of a Pareto front approximation using a performance indicator and such indicators have since been important to measure the performance of metaheuristics on benchmark problems, where the goal is to achieve a good approximation in terms of coverage and closeness to a Pareto front. Indicator-based multiobjective optimization methods use this indicator directly to guide their search, for instance by measuring individual contributions of points to the indicator in the fitness assignment step of evolutionary algorithms. The idea to interpret a population as a vector in a higher-dimensional space of concatenated points, more recently, gave rise to the development of numerical methods (gradient-based, Newton method) for the approximation of the Pareto front using a population (set vector). These methods not only achieve a good closeness to the Pareto front but also good coverage of the dominated subspace. In particular, indicators based on the Lebesgue integral over the dominated subspace (e.g., the hypervolume indicator) and indicators inspired by the Hausdorff metric (e.g. the Delta p metric) have been successfully developed. We show how to compute the Gradients and Hessian matrix of a population of points and how to hybridize these numerical methods successfully with indicator-based metaheuristics. The topic will show an example of the fruitful interplay between research in evolutionary stochastic algorithms and research in numerical deterministic methods, and there are various new perspectives for research on set-oriented optimization.
Website: `https://www.universiteitleiden.nl/en/staffmembers/michael-emmerich`

Short biography

Michael Emmerich is a Germany-born Computer Scientist who currently lives in Finland and in The Netherlands. Since 2016 he is appointed at Associate Professor at Leiden University, The Netherlands, where he leads the Multicriteria Optimization and Decision Analytics Group and since 2019 he is a visiting researcher at Jyvaskyla University, Finland in the Multiobjective and Industrial Optimization Group. He also is Lead AI Scientist at SILO.ai, a provider of AI solutions n the Nordic Countries. He has received his Doctorate in Natural Sciences from the Technical University of Dortmund on the topic of Gaussian Processes for Surrogate-Assisted Multiobjective Design Optimization (2005) under the supervision of Prof. Dr. Ing. H.-P. Schwefel. He also worked as a visiting fellow at the Center for Applied Systems Analysis, ICD e.V. Dorttmund, Institut für Erstarrung unter Schwerelosigkeit e.V. (Aachen), Institute for Fundamental Research of Matter (FOM) Amsterdam, Unversity of the Algarve, IST Lisbon, and Jyvaskyla University. His main contributions are in the field of indicator-based multiobjective optimization, computational geometry algorithms for performance indicator computations, and chemistry and engineering design optimization. He has coordinated four Lorenz Center Workshops and was general chair of three scientific conferences (Global Optimization Workshop 2018 (LeGO), EVOLVE 2013 (Leiden, The Netherlands), and EVOLVE 2015 (Iasi), as well as the organizer and co-initiator of the Modern Machine Learning Technologies (MoMLeT) workshop held annually in Ukraine since 2019. He is general chair of the forthcoming EMO 2023 conference, to be held in Leiden. Together with co-authors he has published five books, 52 journal papers, and more than 120 conference papers, of which five received a best paper award. He has successfully supervised more than 10 Ph.D. students on topics of multiobjective optimization.

Lee Spector

Honor all the things: A lexicase selection manifesto

Amherst College, USA

Abstract

Usually, we care about a lot of things. That is, we have many objectives. Even when we describe a goal in terms of a single objective, that single objective is often an aggregate of many components such as measures of performance on individual tests. The wide-spread, long-standing practice for dealing with such compound objectives is to focus on aggregates, either with a single measure of overall performance or with a collection of measures, each of which aggregates sub-objectives of a specific kind. In this talk I will make the case for an alternative approach that avoids any form of aggregation of objectives or sub-objectives. The lexicase parent selection algorithm, developed for evolutionary computation, is an example of this approach. Rather than aggregating errors on individual tests into an overall "fitness" measure on which selection is then based, lexicase selection uses all of the individual test errors, without aggregation, as the basis for selection. It does this by filtering candidate parents by performance on individual tests, considered sequentially in random order. In doing so, it "honors" all of the individual tests and all possible collections of tests when they appear together at the beginning of test sequences. This can have dramatic effects on search performance, significantly increasing the probability that a solution will be found and decreasing the effort required to find it. I will present an overview of recent work on lexicase selection and speculate about the future of methods that similarly "honor all the things."
Website: `https://lspector.github.io/`

Short biography

Lee Spector is the Class of 1993 Professor of Computer Science at Amherst College, an Adjunct Professor in the College of Information and Computer Sciences at the University of Massachusetts, Amherst, and an Emeritus Professor of Computer Science at Hampshire College. He received a B.A. in Philosophy from Oberlin College and a Ph.D. in Computer Science from the University of Maryland, College Park. His areas of teaching and research include evolutionary computation, quantum computation, and intersections of computer science, cognitive science, and the arts. He is the Editor-in-Chief of the journal Genetic Programming and Evolvable Machines (published by Springer), an Associate Editor for the IEEE Transactions on Evolutionary Learning and Optimization, and a member of the editorial board of Evolutionary Computation (published by MIT Press). He is a member of the ACM SIGEVO executive committee and he was named a Fellow of the International

Society for Genetic and Evolutionary Computation. He has won several other awards and honors, including two gold medals in the Human Competitive Results contest of the Genetic and Evolutionary Computation Conference, and the highest honor bestowed by the National Science Foundation for excellence in both teaching and research, the NSF Director's Award for Distinguished Teaching Scholars.

Laura Cruz Reyes

Where is research going in multi-objective evolutionary optimization incorporating user preferences?

ITCM of TecNM, Mexico

Abstract

A crucial step in solving a multi-objective optimization problem (MOOP) is to identify a set of conflicting solutions in which improving one objective will worsen the performance of the others. This set of optimal solutions in the objective space is the Pareto front. However, obtaining this front does not solve a MOOP because there is usually no single optimal solution. The DM must provide information about his preferences to choose and implement only one solution, the most preferred. In multi-objective metaheuristic approaches, preference information is used to guide the search toward the DM's region of interest and find the best compromise solution. Of interest for this talk is the incorporation of user preferences in multi-objective evolutionary algorithms (EAs). The research area in multi-objective evolutionary optimization began to be recognized in the late 1990s. The first attempt to incorporate preferences in an EA dates back to 1993; in 2000, still very few researchers addressed this issue, and in recent years, research has increased, gaining important advances. The first part of this talk will look at some of these advances, emphasizing the most recent and important ones. In the second part of the talk, a critical analysis of the research developed will be conducted. In the third and last part of the talk, some of the future research challenges will be mentioned.
Website: `http://www.cruz-reyes.com/`

Short biography

Laura Cruz-Reyes received the Ph.D. degree in computer science from the National Center for Research and Technological Development the MS degree in computer science and the MS degree in information systems from the Technological Institute of Monterrey, and a BS degree in Chemical Engineering from the Technology Institute of Ciudad Madero. She is a full-time Professor of Computer Science at the Technological Institute of Ciudad Madero of the National Mexican Institute of Technology. At this institution, she heads the consolidated research group on Intelligent Optimization. She is a member of the Mexican Research System of the government agency CONACYT, with level III, the highest level. She is a member of the Mexican Academy of Computing (AMEXCOMP), the Mexican Society of Operation Research (SMIO), the Mexican Society of Artificial Intelligence (SMIO), and the Mexican Society of Computer Science (SMCC). Her work has focused on modeling and solving complex optimization problems (NP-hard) for environments with many objectives, dynamic conditions, uncertainty, and preferences, supported mainly by artificial intelligence

and operational research disciplines. In this context, her main research interests include metaheuristics, machine learning, fuzzy logic, multicriteria decision, and logistics. She has served as Guest Editor and referee of national and international journals in artificial intelligence and optimization.

Una-May O'Reilly

Modeling Adversarial Dynamics

MIT, USA

Abstract

My interest is in the intelligence of adversaries, particularly how they learn from their conflicts and how their strategic and tactical adaptive behavior can be modeled. I investigate a variety of machine learning methods to model adversarial dynamics. I will describe algorithmic frameworks we have developed for cyber security which draw upon symbolic behavioral descriptions and evolutionary adapation.
Website: `https://alfagroup.csail.mit.edu/unamay`

Short biography

Una-May O'Reilly is founder and leader of the AnyScale Learning For All (ALFA) group at Massachusetts Institute of Technology Computer Science and Artificial Intelligence Laboratory. ALFA focuses on Artificial Adversarial Intelligence through machine learning and evolutionary algorithm lenses. She received the EvoStar Award for Outstanding Achievements in Evolutionary Computation in Europe in 2013. She was a Fellow of the International Society of Genetic and Evolutionary Computation, now ACM SIGEVO, which transferred to special recognition for contributions. She has served as Vice-Chair of ACM SIGEVO. She is the area editor for Data Analytics and Knowledge Discovery for Genetic Programming and Evolvable Machines (Kluwer), and ACM Transactions on Evolutionary Learning and Optimization, and editor for Evolutionary Computation (MIT Press).

Women at NEO

Aim of the session

Women at NEO is a space to promote and intensify research, discussion, and collaboration among the female NEO community. We take time at NEO also for networking and meeting consolidated and young researchers, as well as involved students. One of the goals of W-NEO is to inspire, engage and advice students who are currently working—or planning to work—on optimization subjects lead by the female professors. This section consists of some talks and a short meeting. We will encourage the setting up of specific woman networks with common interests.

It is worth to notice that male researchers/students are also welcome in every part of this session!

Adriana Lara

Marcela Quiroz

Xalapa

NEO 2022 Contributed Talks

Here are all the abstracts of NEO 2022 (NEO X) Workshop. Please, check the schedule (in page 9) in this same notebook to have a list of all the abstracts.

Performance Analysis of Multi-Objective Simulated Annealing Based on Decomposition

Manuel Vargas-Martinez[a], Nelson Rangel-Valdez[b], Eduardo Fernández[c]

[a]División de Estudios de Posgrado e Investigación
Tecnológico Nacional de México / Instituto Tecnológico de Ciudad Madero
Av. 1o. de Mayo esq. Sor Juana Inés de la Cruz s/n Col. Los Mangos C.P.89440, Ciudad Madero, México
D00070864@cdmadero.tecnm.mx

[b]División de Estudios de Posgrado e Investigación
Cátedras CONACyT - Tecnológico Nacional de México / Instituto Tecnológico de Ciudad Madero
Av. 1o. de Mayo esq. Sor Juana Inés de la Cruz s/n Col. Los Mangos C.P.89440, Ciudad Madero, México
nelson.rangel@itcm.edu.mx

[c]Dirección de Investigación y Posgrado
Universidad Autónoma de Coahuila
Blvd. Revolución Oriente 151, Ciudad Universitaria, C.P. 27000, Torreón, Mexico
neddyf171051@gmail.com

Decomposition is an optimization strategy that adds significant diversity to the population in the optimization process. However, increasing convergence in this type of strategy is necessary while maintaining diversity. Traditionally, the decomposition uses information from the neighborhood of each sub-problem to guarantee convergence. Simulated annealing is a metaheuristic strategy that allows for the provision of a balance between exploration and exploitation. This has a type of asymptotic convergence, that is when the algorithm begins it prioritizes exploration, but as time progresses it will prioritize exploitation until it finds an optimal solution. To maintain a balance between diversity and convergence, a simple algorithm based on Multi-Objective Simulated Annealing (MOSA), and Decomposition using vector perturbation operations of Differential Evolution (DE) was designed. The algorithm called MOSA/D takes a multi-objective optimization problem (MOP) and decomposes it into a set of N sub-problem - represented by a set of N weight vectors. Each sub-problem is annealed by L executions at a temperature level T, obtaining new solutions through vector operations. The vector operations of crossover and mutation allow it to achieve convergence without the use of neighborhood information of each sub-problem. Finally, the performance of MOSA/D was analyzed in two variants, one using differential evolution and the other without. Based on DTLZ Benchmark used for evaluation, the result showed a significant difference in favor of the use of differential evolution, observed in the cases of the three, five, and ten objectives.

References

[1] P. Engrand, Multi-objective Optimization Approach Based on Simulated Annealing and its Application to Nuclear Fuel Management, Proceedings of the 5th International Conference on Nuclear Engineering: ICONE-5 , pp. 416-423, 1997.

[2] Q. Zhang, et al., MOEA/D: A Multiobjective Evolutionary Algorithm Based on Decomposition, IEEE Trans. Evol. Comput., vol. 11, no. 6, pp. 712–731, 2007, doi: 10.1109/TEVC.2007.892759

[3] G. Wu , et al., Ensemble of Differential Evolution Variants, Information Sciences, pp. 172–186, 2018, doi: 10.1016/j.ins.2017.09.053

[4] K. Amine, et al., Multiobjective Simulated Annealing: Principles and Algorithm Variants, Advances in Operations Research, vol. 2019, Article ID 8134674, 13 pages, 2019, doi: 10.1155/2019/8134674

Multi-objective optimization through an interactive evolutionary framework with uncertainty

Pablo Hernández-Vicente[a], **Claudia Gómez-Santillán**[a]
Nelson Rangel-Valdez[a], **Fausto Balderas-Jaramillo**[a]

[a]Division of Postgraduate Studies and Research
Instituto Tecnológico de Ciudad Madero
Av. 1o. de Mayo esq. Sor Juana Inés de la Cruz s/n Col. Los Mangos, Cd. Madero, Tamaulipas, México
g15071361@cdmadero.tecnm.mx
claudia.gomez@itcm.edu.mx
nelson.rangel@itcm.edu.mx
fausto.bj@cdmadero.tecnm.mx

Many real-world difficulties require the resolution of multi-objective optimization problems. Solutions that approximate the Pareto front are needed to solve these problems, satisfying the compromise condition of not improving one objective if another is worsened [1]. This approximation to the Pareto front is necessary for the decision maker to choose the final solution.

An alternative to solving multi-objective problems is to use an evolutionary algorithm. This algorithm uses strategies that allow it to generate random solutions and form a population; individuals select from this population evolve through crossbreeding and mutation [2]. Finally, there are new solutions that form the next generation.

The last population generated by these algorithms contains the best solutions; however, this population may be too large, thus complicating the process of selecting the final solution that the decision maker must perform. Therefore, a preference incorporation strategy must be integrated that approximates the interests of the decision maker to facilitate the solution's final choice.

Different parameters are required to model the interests of the decision maker, such as the weights of the objectives to use the previously mentioned preference incorporation strategies. However, these values generally cannot be defined precisely by the decision maker, so ranges or intervals can be used to cover the uncertainty of these values.

The decision maker's preferences can be considered before the execution of the evolutionary algorithm, at the end of the execution, or interactively during the algorithm's execution [3]. This last method is the least studied because the process is more complex and slower than the a priori and a posteriori incorporation due to the intervention of the decision maker.

Therefore, an interactive evolutionary framework has been proposed that uses preference disaggregation analysis and a chat-like interface. Then, through this proposal, the preferences of the decision maker can be efficiently incorporated, the number of tools that integrate this type of incorporation of preferences increases, and it demonstrates that the solutions converge before other types of articulation of preferences. Furthermore, with this proposal, the decision maker can see how the search moves in the solution space thanks to incorporating their preferences, thus facilitating the final choice of the solution.

References

[1] E. Fernández, et al., Preference Incorporation Into Moea/D Using An Outranking Approach With Imprecise Model Parameters, 2021 Swarm and Evolutionary Computation, pp. 2-5, doi: 10.1016/j.swevo.2022.101097

[2] C. Coello, et al., Evolutionary Algorithms for Solving Multi-Objective Problems, 2007 Springer, doi: 10.1007/978-0-387-36797-2

[3] L. Rachmawati, et al., Preference Incorporation in Multi-objective Evolutionary Algorithms: A Survey, 2006 IEEE Congress on Evolutionary Computation, pp. 1-7, doi: 10.1109/CEC.2006.1688414

Coevolutionary algorithm applying outranking method with uncertainty for multi-objective optimization

Lorena Rosas-Solórzano[a], Claudia Gómez-Santillan[a],Laura Cruz-Reyes[a],Nelson Rangel-Valdez[a]

[a]Division of Postgraduate Studies and Research
Instituto Tecnológico de Ciudad Madero
Av. 1o. de Mayo esq. Sor Juana Inés de la Cruz s/n Col. Los Mangos, Cd. Madero, Tamaulipas, México.
lorenarocio52@gmail.com
cggs71@hotmail.com
cruzreyeslaura@gmail.com
rangel.nelson1980@gmail.com

The strategies for solving real-life multi-objective optimization problems have evolved widely, but challenges such as algorithmic design applied to large-scale problems still exist.

The scalability problems are one of the challenges in open issues that are receiving attention from the research community because they are not correctly modeled, and the algorithms would not find the optimal results. Martí, L. [1] comments that traditional strategies, for example, those that use Pareto ordering, when increasing the number of objectives, do not achieve good performance because the algorithm does not work well by increasing the number of objectives and variables.

So, at present, it is proposing to address scalability problems through the use of co-evolutionary algorithms, cellular, among others, without affecting the effectiveness of the problem [4, 5]. To treat scalability, co-evolutionary algorithms divide the problem into subproblems. Each one evolves in its population, adapting to the environment through repetitions applied in an evolutionary algorithm (EA), having reciprocal evolutionary changes between the participants. In the end, the representatives of each population come together [6]. Some works, such as those of Lohn, J. [7] and Antonio [4,5], present co-evolutionary algorithms to solve multi-objective problems. In addition to the scalability problems inherent in this work due to the characteristics of the problems that are solved, other variants that include some decision-making model and uncertainty will also be addressed.

Fernández [8] and Balderas [9] have worked with multi-objective problems and have applied them to overcome relationships and intervals; that is, they compare the alternatives in pairs for each criterion, finding the strength to prefer one over the other [10]. These problems also include uncertainty, which is present at the thresholds of the overcoming model and is represented by interval mathematics to solve multi-objective problems, where there is no precise knowledge or there may be variabilities in the data.

In the literature, works in multi-objective optimization apply uncertainty [8, 11, 12] and co-evolution [13, 14] separately. Still, no works have been found where they combine the two strategies in one. Only algorithm, also another contribution is to include an overclassification model with the preferences of the decision maker; for this reason, it is proposed to incorporate these strategies in a co-evolution algorithm that makes use of the outranking method to solve multi-objective optimization problems with uncertainty, where not only will scalability be addressed in the number of decision variables, but also the number of objective functions.

References

[1] Martí, L. , Scalable Multi-Objective Optimization, 2011, Colmenarejo, Universidad Carlos III de Madrid. Escuela Politecnica superior

[2] Tse, Guan Tan, and Teo, Jason Tze Wi and Hui, Keng Lau , Performance scalability of a cooperative coevolution Multiobjective Evolutionary Algorithm, 2007 International Conference on Computational Intelligence and Security (CIS'07), Harbin, Heilongjiang, China

[3] Carlos M. Fonseca1, Kathrin Klamroth , Günter Rudolph , and Margaret M. Wiecek., Scalability in Multiobjective Optimization, Seminar January 12–17, 2020, Dagstuhl Reports, Vol 10. Issue 1, pp. 52-129

[4] Antonio, L. Coello, A. , Use of Cooperative Coevolution for Solving Large Scale Multiobjective Optimization Problems, 2013, IEEE Congress on Evolutionary Computation, June 20-23, Cancún, México

[5] Antonio, L., Coello, C., Morales, M., Brambila, S., González, J. and Tapia, G, Coevolutionary Operations for Large Scale Multi-Objective Optimization, 2020 IEEE Congress on Evolutionary Computation (CEC), pp. 1-8, doi: 10.1109/CEC48606.2020.9185846

[6] Potter, M. and De Jong, K., A. Cooperative Coevolution: An Architecture for Evolving Coadapted Subcomponents, 2000 Evolutionary Computation 8(1), pp. 1-29.

[7] Lohn, J., Kraus, W. Haith, G., Comparing a Coevolutionary Genetic Algorithm for Multiobjective optimization, 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No.02TH8600), pp. 1157-1162 vol.2, doi: 10.1109/CEC.2002.1004406

[8] Fernández, E. , Interval-based extensions of two outranking methods for multi-criteria ordinal classification, 2019 The International Journal of Management Science.

[9] Balderas, F., Fernández, E., Gómez, C., Cruz, L. y Rangel, N. , Metaheuristic Robust Optimization of Project Portfolios using an Interval-Based Model of Imprecisions, 2016, International Journal of combinatorial optimization problems and informatics. Vol 7, No. 3 pp. 101-118, ISSN: 2007-1558.

[10] Tscheikner-Gratl, F., Egger, P., Raunch, W. y Kleidorfer, M., Comparison of Multi-Criteria Decision Support Methods for Integrated Rehabilitation Prioritization, 2017, doi: 10.3390/w9020068

[11] Balderas, F., Fernandez, E., Gomez, C., Valdez, N. Cruz, L., An interval-based approach for evolutionary multi-objective optimization of project portfolios, 2019 International Journal of Information Technology Decision Making. Vol 18, No 14, pp. 1317-1358.

[12] Fernández, E., Gómez, C., Rangel, N. Cruz, L. , Group decision and negotiation. Springer Fonseca C., Klamroth, K. Rudolph and Wiecek, M. Reformulating preferences into constraints for evolutionary multi- and many-objective optimization., 2020 Information Sciences Information Sciences 541 (2020).

[13] Wang, R., Purshouse, R. C., Fleming, P. J., Preference-Inspired Coevolutionary Algorithms for Many-Objective Optimization, 2013 IEEE Transactions on Evolutionary Computation, 17(4), pp. 474–494, doi: 10.1109/tevc.2012.2204264

[14] Purshouse, R. C., Jalbă, C., Fleming, P. J. , Preference-Driven Co-evolutionary Algorithms Show Promise for Many-Objective Optimization. Evolutionary Multi-Criterion Optimization, 2011, pp. 136–150, doi: 10.1007/978-3-642-19893-9_10

Hybrid Harmonic Search Algorithm with Improved Opposition-Based Learning (HHS-IOBL) in Benchmark Functions

José Alfredo Brambila-Hernández[1], Eduardo Villegas-Huerta[1*], Héctor Joaquín Fraire[1*], Miguel Ángel García-Morales[1], Armando Becerra de Angel [1]

[1] Tecnológico Nacional de México/Instituto Tecnológico de Ciudad Madero, Posgrado e Investigación, Madero, Tamaulipas,
México

jabrambila@gmail.com (J.A.B-H), automatas2002@yahoo.com.mx (H.J.F-H),soporteclusterlanti@gmail.com (M.A.G-M), eduardo.villegas.h@hotmail.com (E.V-H), armando.ba@cdmadero.tecnm.mx (A-B-A)

Special sessions: Decision Making and Intelligent Agents

Abstract. This paper proposes a hybrid harmonic search algorithm that incorporates a method of reinitializing harmonies memory using a particle swarm optimization algorithm with an improved opposition-based learning method (IOBL) to solve continuous optimization problems. This method allows the algorithm to obtain better results by increasing the search space of the solutions. This approach has been validated by comparing the performance of the proposed algorithm with that of a state-of-the-art harmonic search algorithm, solving nine standard mathematical functions, and applying the Wilcoxon parametric test at a 5% significance level. The state-of-the-art algorithm uses an opposition-based improvement method (IOBL). Computational experiments show that the proposed algorithm outperforms the state-of-the-art algorithm. In quality, it is better in eight out of nine instances, and efficiency is better in six out of nine instances.

Keywords: Harmonic Search, Improved Learning, Opposition-based, Hybrid Algorithm.

1. Introduction

Currently, society requires maximum benefits at minimum costs. In order to achieve this, optimization techniques are generally used. However, many real-world optimization problems are considered high computational complexity and are called NP-hard. Moreover, these problems have the characteristic that exact solution methods cannot obtain optimal solutions in reasonable times. Using metaheuristic methodologies is considered a good alternative that offers satisfactory solutions for the user in a reasonable time. Harmonic search algorithms constitute a metaheuristic

methodology for solving continuous optimization problems proposed by Zong Woo Geem [1].

The process of musical improvisation inspires the methodology of harmonic search. In this, a predefined number of musicians try to tune the tone of their instruments until they achieve a pleasant harmony. In nature, *harmony* is a relationship between several sound waves with different frequencies. Therefore, the quality of improvised harmony is determined by aesthetic estimation. In order to improve aesthetic valuation and find the best harmony, musicians perform multiple rehearsals [2].

Harmonic search algorithms are currently considered a competitive alternative to solve a large number of optimization problems that have several advantages over other metaheuristics that are available in the state-of-the-art. For example, they only require adjusting a relatively small number of parameters [3–5].

In [6], a harmonic search algorithm using an improved OBL mechanism (IOBL) is proposed. This improved version uses randomness to create a new possible solution and improve the convergence process of such an algorithm. In addition, the IOBL mechanism is used in the upgrade process.

The main contribution is the use of a mechanism of a reinitialization of the memory of harmonies using a particle swarm optimization algorithm to improve the quality of the solutions produced by the harmonic search algorithm, in addition to using the improved OBL technique (IOBL) in the process of a reinitialization of the memory of harmonies.

Contribution of the authors: Conceptualization: H.J.F.-H.; Methodology: E.V-H., Research: E.V-H.; Software: J.A.B-H., M.A.G-M.; Formal analysis: H.J.F.-H.; Writing, proofreading, and editing: H.J.F-H, J.A.B-H, M.A.G-M. All authors read and agree to the publication of the article.

Acknowledgments: The authors thank CONACYT for supporting the projects with number A1-S-11012 of the Call for Basic Scientific Research 2017–2018 and project number 12397 of the Support Program for Scientific, Technological and Innovation Activities (PAACTI) in order to participate in the Call 2020-1 Support for Scientific Research Projects, Technological Development and Innovation in Health in the face of the Contingency by COVID-19. Alfredo Brambila, Miguel Angel García Morales and Eduardo Villegas Huerta thank CONACYT for the support 760308, 731279 and001818 respectively. Héctor Fraire thanks the National Technological of Mexico General Directorate for supporting the research project 10362.21-P.

Conflicts of interest: Authors declare that they have no conflict of interest.

References

1. Geem, Z. W., Kim J. H., and Loganathan G. V., A new heuristic optimization algorithm: harmony search, simulation, vol. 76, pp. 60-68. (2001)
2. Askarzadeh, A., Esmat R., Harmony Search Algorithm: Basic Concepts and Engineering Applications, Recent Developments in Intelligent Nature-Inspired Computing. edited by Patnaik, Srikanta, pp. 1-36. Hershey, PA: IGI Global. (2017)
3. Fesanghary, M.; Mahdavi, M.; Alizadeh, Y. Hybridizing harmony search algorithm with sequential quadratic programming for engineering optimization problems. Comput. Methods Appl. Mech. Eng. 197, pp. 3080–3091. (2008)
4. Geem, Z.W.; Sim, K.-B. Parameter-setting-free harmony search algorithm. Appl. Math. Comput. 217, pp. 3881–3889. (2010)
5. Vasebi, A.; Fesanghary, M., Bathaee, S.M.T. Combined heat and power economic dispatch by harmony search algorithm. Int. J. Electr. Power Energy Syst. 29, pp. 713–719. (2007)
6. Alomoush, A. A., Alsewari, A. R. A., Zamli, K. Z., Alrosan, A., Alomoush, W., & Alissa, K. (2021). Enhancing three variants of harmony search algorithm for continuous optimization problems. International Journal of Electrical & Computer Engineering (2088-8708), 11(3).

COVID-19 data analysis with a Multi-objective Evolutionary Algorithm for Causal Association Rules Mining

Santiago Sinisterra-Sierra[a], Salvador Godoy-Calderón[b], Miriam Pescador-Rojas[c]

Centro de Investigación en Computación
Instituto Politécnico Nacional
Ciudad de México, México
[a] ssinisterras2020@cic.ipn.mx
[b] sgodoyc@cic.ipn.mx
Escuela Superior de Cómputo
Instituto Politécnico Nacional
Ciudad de México, México
[c]mpescadorr@ipn.mx

Association rule mining plays a crucial role in the medical area in discovering interesting relationships among the attributes of a data set. Traditional association rule mining algorithms such as Apriori, FP growth, or Eclat require considerable computation resources and generate large volumes of rules. Moreover, these techniques depend on user-defined thresholds which can inadvertently cause the algorithm to omit some interesting rules. In order to solve such challenges, we propose an evolutionary multi-objective algorithm based on NSGA-II to guide the mining process in a data set composed of 15.5 million records with official data describing the COVID-19 pandemic in Mexico. We tested different scenarios optimizing classical and causal estimation measures on 4 waves defined as the periods of time where the number of infected people increased. The proposed contributions generate, recombine and evaluate patterns, focusing on recovering promising high-quality rules with actionable cause-effect relationships among attributes to identify which groups are more susceptible to disease or what combinations of conditions are necessary to receive certain types of medical care.

Relationship between the development of COVID-19 and poverty in the municipalities of Mexico

Christian Lizbeth Noguez-Moreno[a], Edwin Montes-Orozco [b], Ángel David Téllez-Macías [a], Gilberto Sinuhé Torres-Cockrell[c], Roman Anselmo Mora-Gutiérrez[d] and Miguel Ángel Fernández-Romero[e]

[a] Departamento de Tecnologías de la Información. UAM-Cuajimalpa [b] Departamento de Matemáticas Aplicadas y Sistemas. UAM-Cuajimalpa [c] Posgrado en Ciencias y Tecnologías de la Información. UAM-Iztapalapa [d] Departamento de Sistemas. UAM-Azcapotzalco [e] Independent consultor

Abstract

The present work presents an analysis and characterization of the relationship between poverty and the development of Covid-19. For this, we developed a coupled complex network model, in which the level of marginalization and the number of cases and deaths from Covid in each of the municipalities of Mexico were considered. Afterward, we used a variant bi-objective of the Semi-coloring optimization problem to identify communities on the network. We can detect three groups in the structure of the coupled complex network based on numerical results.

1 Introduction

The Covid-19 pandemic severely affected social behavior; its relationship with poverty has been the topic of several articles such as [1, 2, 3, 4]
[2] mentioned that socio-economic factors impact all facets of human functioning, including health-related quality of life. Also, [4] established that the poverty is a factor in the risk for death by COVID-19.
Based on previous information, a model complex network for Mexico is developed and analized through a semi-coloring optimization problem (SCOP) in this work.

2 Methodology

First, the information about cases and death by municipalities of Mexico in 2020 was analyzed, as this information a complex network model is bult. After, the poverty information on municipalities of Mexico (2019-2020) was used to build a complex network model. Next, both models were coupled.
The network structures were studied as the variant of SCOP presented in equation 1. In general terms, this variant of SCOP can be defined as given an undirected network $R : (V, E)$ with weights on the edges; given the complementary lattice $\bar{R} : (V, \bar{E})$ and a number k of colors. A valid partial coloring of the complementary network $V : C_1 \cup C_2 \cup C_k \cup S$ (C_i is formed by the nodes colored by the ith value, while the uncolored nodes form S) is found, such that the sum of the edges of the vertices with the same color in the original network (h) is maximized, and the sum of the edges of the uncolored vertices g is minimized

$$
\begin{array}{rcl}
\max(z_1) & : & h \\
\min(z_2) & : & g \\
\min(\sum w_{i,j}e_{i,j,1}, \sum w_{i,j}e_{i,j,2}, \ldots \sum w_{i,j}e_{i,j,k}) & = & h \\
\sum w_{i,j}e_{i,j,s} * |S| & = & g \\
v_{i,k}\bar{e}_{i,j,k} + v_{j,k}\bar{e}_{i,j,k} & \leq & 1 \\
C_1 \cup C_2 \cup C_k \cup S & = & V \\
C_i \cap C_j & = & \emptyset \\
C_i \cap S & = & \emptyset \\
\sum v_{i,k} & \geq & 1 \\
v_{i,k} \in [0,1] \text{ and it is integer } \forall i = 1, 2, ..., n \; \forall k = 1, ..., K & & \\
\text{each subgraphs } C_k \text{ is connected} \forall k = 1, ..., K & &
\end{array}
\tag{1}
$$

3 Results

Based on communities identified the structure of complex networks, we can say there is a relationship between poverty and the developing of Covid-19 in Mexico

References

[1] Ahmad, K., Erqou, S., Shah, N., Nazir, U., Morrison, A. R., Choudhary, G., & Wu, W. C. (2020). Association of poor housing conditions with COVID-19 incidence and mortality across US counties. PloS one, 15(11), e0241327.

[2] Hawkins, R. B., Charles, E. J., & Mehaffey, J. H. (2020). Socio-economic status and COVID-19–related cases and fatalities. Public health, 189, 129-134.

[3] Kiaghadi, A., Rifai, H. S., & Liaw, W. (2020). Assessing COVID-19 risk, vulnerability and infection prevalence in communities. PLoS One, 15(10), e0241166.

[4] Gerayeli, F. V., Milne, S., Cheung, C., Li, X., Yang, C. W. T., Tam, A., ... & Sin, D. D. (2021). COPD and the risk of poor outcomes in COVID-19: A systematic review and meta-analysis. EClinicalMedicine, 33, 100789.

Binaural and multi-tone generator to characterize nonlinear systems for healthcare and device manufacturing applications

Roberto Herrera Charles[a], **Teodoro Álvarez Sánchez**[a], **Jesus A. Álvarez Cedillo**[b]

[a]Department Rresearch
Instituto Politécnico Nacional
Ave. Instituto Politécnico Nacional No. 1310 Colonia Nueva Tijuana
robccharles@ipn, talvarezs@ipn.mx

[b]Department SEPI
Instituto Politécnico Nacional
Ave. Te 950, Granjas México, Ciudad de México, México, C.P. 22435
jalvarez@ipn.mx

In an audio signal as binaural or multitone is defined as a desirable test signal for fast frequency response measurements, as well as to be able to evaluate the non-linearity of a system under test (DUT). Measurements of conventional multitone signals are challenging, because the number of intermodulation tones grows rapidly as the number of tones. This makes it extremely difficult to separate these harmonics produced by intermodulation. The first contribution of this work is to perform the measurement of intermodulation distortion that stimulates the device under test (DUT), through multiple sinusoidal signals, which are closer to a real situation of a transmission of audio signals, under characteristics of non-linear transfer, the second contribution is to generate new harmonics and inter-modulated frequencies in the device (DUT). This work shows the design of the binaural and/or multitone audio tone generation system to measure the response to audible frequencies of nonlinear systems which have various applications such as health and the characterization of audio devices in productions.

References

[1] TIA/EIA-810-Al., Transmission Requirements for Narrowband Voice over IP and Voice over PCM Digital Wireline Telephones, Telecommunication Industry Association, December 2000,

[2] NTIl, Application Note A2 Rapit Test. Comparing Conventional vs. Multi-tone Testing,

[3] S. Temme and P. Brunetl, A New Method for Measuring Distortion using a Multi-tone Stimulus and Non-Coherence, AES, Presented at the 121st Convention 2006 October 5–8 San Francisco, CA, USA,

A comparison of the behavior of the traditional education system and the remote teaching emerging project implemented by Autonomous Metropolitan University into the emergence of COVID-19

Oswaldo Sánchez Andrade [a], Roman A. Mora-Gutiérrez [b], Edwin Montes-Orozco[c], Sergio G. de los Cobos Silva[d], Eric A. Rincón-García[d], Pedro Lara-Velázquez [d], Miguel A. Gutiérrez Andrade [d], Jorge Juárez Gómez[e].

[a] Posgrado en Diseño y vizualización de la información. UAM-Azcapotzalco [b] Departamento de sistemas. UAM-Azcapotzalco [c] Departamento de Matemáticas Aplicadas y Sistemas. UAM-Cuajimalpa [d] Departamento de ingeniería electríca. UAM-Iztapalapa [e] Departamento de Química. UAM-Iztapalapa

Abstract

In this paper we present a comparison of the behaviors of two teaching-learning systems. This comparison uses a complex network of sets and a variant of the vertex separator problem (VSP) solved by a genetic algorithm. The numerical results show that there is a difference between the two teaching-learning systems, which have different advantages and disadvantages in the face of the new educational models emerging during the COVID-19 pandemic.

1 Introduction

The study of communities in the complex network has been an essential topic in the last decade. In [1] mentioned that the problems of identification and determination of communities are fundamental to the analysis of structure in complex networks. In general terms, a community is a subset of firmly related nodes [2].
On the other hand, the health emergency unleashed by covid-19 involved several changes in education systems worldwide. The Autonomous Metropolitan University implemented the remote emergent learning program in this health emergency. The structure of complex networks (traditional and emergent education systems) is analyzed, characterized, and compared.

2 Methodology

First, the traditional and emergent education systems were modeled as complex networks. In those models, a node represents a teaching-learning (UEA) unit or a professor; it has associated with a vector of the average grades vector. Two UEA have a link if they are similar; a professor and a UEA have a link if the i-th professor has imparted the j-th UEA; two professors have a link if both are similar. Each link is associated with a weight that is represented by graded similarities.
Next, We identified two communities in each complex network, for which the variant of VSP, shown in equation 1, was solved. In general terms, this VSP variant can be defined as: Given a graph $G = (V, E)$, where each edge is associated with a similarity value $w_{i,j}$ and each node has a vector n-dimensional y_i^n. It is desired to determine the non-empty subsets A, B and C of V; such that $A \cap B = \emptyset$, $A \cap C = \emptyset$ $B \cap C = \emptyset$ and $A \cup B \cup C = V$; besides having a certainty of 95% that the averages of the vectors $y_{i,A}^n$ and $y_{i,B}^n$ are different and a certainty of 95% that the weight average of the edges in the subsets A and B are similar.

$$
\begin{gathered}
\min \sum_{i=1}^{|V|} \sum_{j=1\ j\neq i}^{|V|} x_{i,C} x_{j,C} * w_{i,j} \\
\bar{y_A} = \frac{\sum_i^{|V|} x_{i,A}}{|A|} * y_i^n \\
\bar{y_B} = \frac{\sum_i^{|V|} x_{i,B}}{|B|} * y_i^n \\
s_A^{2n} = \frac{\sum_i^{|V|} (x_{i,A})}{|A|} * (y_i^n - \bar{y_A})^2 \\
s_B^{2n} = \frac{\sum_i^{|V|} (x_{i,B})}{|B|} * (y_i^n - \bar{y_B})^2 \\
s^n = \sqrt{(\frac{(|B|-1)*s_B^{2n} + (|A|-1)*s_A^{2n}}{(|B|-1)+(|A|-1)})} \\
|\frac{\bar{y_A^k} - \bar{y_B^k}}{s^k \sqrt{1/|B| + 1/|A|}}| \geq 1.96 \ \forall k = 1, 2, \ldots, n \\
\bar{w_A} = \sum_{i=1}^{|V|} \sum_{j=1\ j\neq i}^{|V|} \frac{x_{i,A} x_{j,A} w_{i,j}}{|A|} \\
\bar{w_B} = \sum_{i=1}^{|V|} \sum_{j=1\ j\neq i}^{|V|} \frac{x_{i,B} x_{j,B} w_{i,j}}{|B|} \\
sw_A^2 = \sum_{i=1}^{|V|} \sum_{j=1\ j\neq i}^{|V|} \frac{(x_{i,A} x_{j,A} w_{i,j} - \bar{w_A})^2}{|A|} \\
sw_B^2 = \sum_{i=1}^{|V|} \sum_{j=1\ j\neq i}^{|V|} \frac{(x_{i,B} x_{j,B} w_{i,j} - \bar{y_B})^2}{|B|} \\
sw = \sqrt{(\frac{(|B|-1)*sw_B^2 + (|A|-1)*sw_A^2}{(|B|-1)+(|A|-1)})} \\
|\frac{\bar{w_A} - \bar{w_B}}{s \sqrt{1/|B| + 1/|A|}}| \leq 1.96 \\
x_{i,A} + x_{i,B} + x_{i,C} = 1 \ \forall i = 1, 2, \ldots, |V| \\
|A| \geq 1 \\
|B| \geq 1 \\
|C| \geq 1 \\
|A| = \sum_{i=1}^{|V|} x_{i,A} \\
|B| = \sum_{i=1}^{|V|} x_{i,B} \\
|C| = \sum_{i=1}^{|V|} x_{i,C} \\
|A| + |B| + |C| = |V| \\
x_{i,k} \in [0, 1] \text{ and it is integer } \ \forall i = 1, 2, ..., n \ \forall k = A, B, C \\
\text{subgraphs A and B are connected.}
\end{gathered}
\tag{1}
$$

3 Results

Based on communities identified about the structure of complex networks, we can say two learning-teaching are different behavior. Communities a look at each network have different behavior.

References

[1] Chakraborty, T., Srinivasan, S., Ganguly, N. et al. Constant Communities in Complex Networks. Sci Rep 3, 1825 (2013). https://doi.org/10.1038/srep01825

[2] Balakrishnan, H., & Deo, N. (2006, March). Discovering communities in complex networks. In Proceedings of the 44th annual Southeast regional conference (pp. 280-285).

A genetic algorithm for predicate discovery with compensating fuzzy logic

José Fernando Padrón-Tristán[a], **Laura Cruz-Reyes**[b], **Rafael Alejandro Espín-Andrade**[c], **Carlos Eric Llorente-Peralta**[a]

[a]Division of Postgraduate Studies and Research
Instituto Tecnológico de Tijuana
Tijuana, Baja California
jfpt76@gmail.com
ceric.lop@gmail.com

[b]Division of Postgraduate Studies and Research
Instituto Tecnológico de Ciudad Madero
Ciudad Madero, Tamaulipas
lauracruzreyes@itcm.edu.mx

[b]Facultad de Contaduría y Administración. Dirección de Postgrado
Universidad Autónoma de Coahuila
Torreón, Coahuila
rafaelalejandroespinandrade@gmail.com

Compensatory Fuzzy Logic (CFL) starts from the idea that the way in which human thought is built is not through numbers, but on linguistic labels and considers that the certainty of a proposition is a matter of degree [1]. Thus, the most attractive features of fuzzy logic are its flexibility, its tolerance for imprecision, its ability to model nonlinear problems, and its basis in natural language. The CFL is a multivalent system that extends the traditional axiomatics of this type of system to achieve semantically better behavior than classical systems [1].

According to Serov et al. (2019), a fuzzy predicate is a function defined in fuzzy variables whose range of values is in the interval (0,1). The result of this function, called the truth value, is obtained with the composition of fuzzy variables (known as individual fuzzy predicates) and fuzzy logic operators [2].

The CFL is formed by a quartet of continuous operators conjunction (c), disjunction (d), fuzzy strict order (o) and negation (n) which satisfy a group of axioms, among which are those of compensation and veto [3].

In the present research work, a genetic algorithm (GA) is developed that generates predicates that comply with a defined structure in a search configuration that includes novel elements.

GA are search algorithms based on concepts of natural selection and natural genetics. GAs are highly parallel, mathematical, and adaptive search procedures (ie, problem-solving methods) loosely based on the processes of natural genetics and Darwinian survival of the fittest [4].

These algorithms apply genetically inspired operators to populations of possible solutions iteratively, creating new individuals in the population while searching for an optimal (or near-optimal) solution to the problem at hand [4].

The developed GA includes elements, among which the use of deductive logical structures and interpretability indices stands out, to guide the discovery of predicates and which are listed below:

• a method of controlling the diversity of fuzzy predicates,

• a method to control the interpretability of fuzzy predicates that is integrated into the predicate construction process and,

• a method of building fuzzy predicates using non-recursive genetic programming.

This proposal is a work in progress, in which initially promising results have been found.

References

[1] Perez-Pueyo R., (2005)., Procesado y optimización de espectros raman mediante técnicas de lógica difusa: aplicación a la identificación de materiales pictóricos, Universitat Politecnica de Catalunya. Departament de Teoria del Senyali

[2] Serov, V. V., Sokolov, I. V., and Budnik, A. A. (2019), Applied calculus of fuzzy predicates for the formalization of knowledge, In IOP Conference Series: Materials Science and Engineering (Vol. 537, No. 4, p. 042043). IOP Publishing, `https://iopscience.iop.org/article/10.1088/1757-899X/537/4/042043/pdf`

[3] Espin-Andrade, R. A., Gonzalez, E., Pedrycz, W., and Fernandez, E. (2016), An interpretable logical theory: the case of compensatory fuzzy logic. International Journal of Computational Intelligence Systems, 9(4), 612-626

[4] Mathew, T. V. (2012), Genetic algorithm. Report submitted at IIT Bombay.

A cooperative coevolutionary genetic approach to solve packing problems

Jordan Michelt Aran Perez[a], Laura Cruz Reyes[a] Bernabé Dorronsoro[b], Héctor Fraire[a], Nelson Rangel Valdeza[a], Claudia Gómez Santillán[a] Marcela Quiroz Castellanos[c]

[a]Division of Postgraduate Studies and Research
Instituto Tecnológico de Ciudad Madero
Av. 1o. de Mayo esq. Sor Juana Inés de la Cruz s/n
Col. Los Mangos, Cd. Madero, Tamaulipas, México
g16070192@cdmadero.tecnm.mx
lauracruzreyes@itcm.edu.mx
hector.fraire2014@gmail.com
cggs71@hotmail.com
rangel.nelson1980@gmail.com

[b]Computer Science Engineering Department
University of Cadiz
Avenida de la Universidad, 10
11519 Puerto Real, Cadiz, Spain
bernabe.dorronsoro@uca.es

[c]Artificial Intelligence Research Center
Universidad Veracruzana
Calle Salvador Díaz Mirón 35, Zona Universitaria,
Xalapa-Enríquez, Veracruz, México.
maquiroz@uv.mx

One important type of computational problem is the grouping problems, where a set of elements is divided into a collection of subsets. One of these problems is the One-Dimensional Bin Packing Problem (1D-BPP), which is a classic optimization problem known for its applicability and complexity.1D-BPP belongs to a special class of problems called NP-hard, in which, given a set of variable size of elements, we pursue to accommodate them within fixed size containers, seeking to optimize the number of containers to use, that is, using the smallest number of containers to place the largest number of items possible [1].

The 1D-BPP is one of the most fundamental problems in combinatorial optimization and has been widely studied for decades. As one of the most representative problems of the constant grouping problems, new sets of instances are proposed and also new algorithms that try to solve it. The 1D-BPP has many applications in real life, hence its relevance and also serves as a starting point to understand other more complex grouping problems.

There are several metaheuristic strategies proposed to solve the 1D-BPP. However, most of these approaches have focused on the resolution of specific instances and have not considered the generality in order to increase the capacity of solution in problems or instances [2]. One of the strategies that helps the generality is the Coevolution.

The objective of this work is to introduce a Cooperative Coevolutionary Genetic Algorithm (CCGA) to solve the 1D-BPP problem, by creating subspecies focused on improving the population in different aspects. For this, we propose a coding scheme and genetic operators adapted to a coevolutionary approach that allows the division of species. The proposal is based on the state-of-the art algorithm called the Grouping Genetic Algorithm with Controlled Gene Transmission (GGA-CGT) [3].

As a result, a better understanding of the impact of including coevolutionary strategies in the resolution of mono-objective grouping problems is presented. Finally, a general overview of CCGA and its performance

is discussed, including suggestions for addressing other grouping problems with similar characteristics to 1DBPP.

References

[1] Garey, M. R., & Johnson, D. S., Computers and intractability: a guide to the theory of NPcompleteness (series of books in the mathematical sciences). 1st ed., W. H. Freeman, editor. New York: W. H. Freeman and Company (1979).

[2] Pillay, N., Evolving hyper-heuristics for the uncapacitated examination timetabling problem. Journal of the Operational Research Society, 63(1), 47-58 (2012).

[3] Quiroz-Castellanos M., Cruz-Reyes L., Torres-Jimenez J., Gómez S. C., F. Huacuja H., & Alvim A., A grouping genetic algorithm with controlled gene transmission for the bin packing problem. Computational Operation Research pp.52–64 (2015).

Study of Methods for One-dimensional Bin Packing Problem

Jessica González-San Martín[a], **Laura Cruz-Reyes**[a]
Héctor Fraire[a], **Bernabé Dorronsoro**[b]

[a]Division of Postgraduate Studies and Research
Tecnológico Nacional de México, Instituto Tecnológico de Ciudad Madero
Av. 1o. de Mayo esq. Sor Juana Inés de la Cruz s/n
Col. Los Mangos, Cd. Madero, Tamulipas, México
jessica_san_martin@hotmail.com, lauracruzreyes@itcm.edu.mx
hector.fraire2014@gmail.com

[b]Computer Science Engineering Department
University of Cadiz
Avenida de la Universidad, 10 C.P.11519 Puerto Real,
Cadiz, Spain
bernabe.dorronsoro@uca.es

The Bin Packing Problem (BPP) is a classic optimization problem that is known for its applicability and complexity, which belongs to a particular class of problems called NP-hard, in which, given a set of items of variable size, we search to accommodate them inside fixed size containers, seeking to optimize the number of containers to be used, that is, using the least number of containers to place the largest number of items possible. For this problem, there are different variants with respect to the dimensions of objects (1D-BPP, 2D-BPP and 3D-BPP), number of objectives to satisfied (Multiobjective) and changes on time (Dynamic). In this study we will focus only on the One-dimensional Bin Packing Problem (1D-BPP), because it is the base problem for the multidimensional, multiobjective and dynamic variants and has been preserved as a current study problem due to the various applications that it offers. In the current state of the art for 1D-BPP, there are different algorithms, mainly heuristics for solving the problem, however, there is no heuristic or metaheuristic algorithm capable of finding the optimal solution for all possible instances of a problem of this type despite the scientific community's efforts.

This work presents a study of methods and strategies that have been used to address 1D-BPP in the last two decades, in order to identify the most promising ones regarding algorithmic performance. The main objective is that this study can help both researchers and professionals interested in using specific components or techniques that help improve the behavior of an algorithm to solve this problem.

Among the strategies implemented in recent years we find the hybridization of metaheuristic algorithms with machine learning, with reinforcement learning (RL) being the most used, which is an excellent alternative to dynamically adapt the search for these heuristics by training an agent in a supervised or self-supervised manner [1]. A reinforcement learning agent tries to learn the behavior through trial-and-error interactions in a dynamic and uncertain environment. The agent does not know what actions to take, and it must automatically discover which actions produce the maximum benefit. In addition, it must be able to adapt to changes that occur in the environment. On each interaction, the agent receives an indication of its current state and selects an action. The action changes state and the agent receives a reinforcement or reward signal. The agent's goal is to find a policy that associates states with actions, maximizing the long-term accumulated reward.

Currently, we see the field of RL for combinatorial optimization (CO) problems as a promising direction for this research line because of the effectiveness in terms of the solution quality, the capacity to outperform the existing algorithms, and huge running time gains compared to the classical heuristic approaches.

We present a grouping genetic algorithm as case of study. In the proposed version, a new strategy was implemented incorporating reinforcement learning to one of the genetic operators. Preliminary results of our version showed an increase in its performance compared to the original genetic algorithm [2]. According to the results obtained in the experimentation, the proposed version improves the performance of the original algorithm by 6.23%, solving 80 more instances and reducing the execution time. With this work, it was possible to identify some promising strategies that could be addressed in the future for 1D-BPP and obtain a greater impact on performance if are combined or implemented in conjunction with other methods.

References

[1] Mazyavkina, N., Sviridov, S., Ivanov, S., & Burnaev, E., Reinforcement learning for combinatorial optimization: A survey. Computers & Operations Research, 134, 105400. 2021

[2] Quiroz-Castellanos M., Cruz-Reyes L., Torres-Jimenez J., Gómez S. C., F. Huacuja H., & Alvim A., A grouping genetic algorithm with controlled gene transmission for the bin packing problem. Computational & Operations Research pp.52–64. 2015

A new high performance crossover operator for the Grouping Genetic Algorithm with Controlled Gene Transsmision

Stephanie Amador-Larrea [a], Marcela Quiroz-Castellanos [a], Guillermo-de-Jesús Hoyos-Rivera [a]

[a]Artificial Intelligence Research Institute
Universidad Veracruzana
Campus Sur, Calle Paseo Lote II, Sección Segunda N° 112, Nuevo Xalapa, Xalapa, Ver., México 91097
stephanieamadorlarrea@gmail.com, maquiroz@uv.mx, ghoyos@uv.mx

The one-dimensional Bin Packing Problem (1D-BPP) has been a problem widely studied over the last four decades, because of its multiple applications in industry, logistics, transportation and others. Due to the complexity of finding an optimal solution in a short period of time, this problem has been approached using different techniques, highlighting the performance of hybrid algorithms and heuristics. Among the most outstanding state-of-the-art proposals is the GGA-CGT [1], which has outperformed the best algorithms in high-difficulty benchmarks. Nevertheless, experimental studies revealed that, for a significant number of new 1D-BPP instances, the GGA-CGT predeterminated crossover operator did not seem to lead to better solutions. For this reason, an experimental study of group oriented crossover operators was performed within the state of the art, using 1D-BPP instances with different characteristics and a high level of difficulty, with the objective of studying the impact that each operator can have on the final performance of the GGA-CGT [2]. In order to identify the characteristics that a crossover operator should have to contribute significantly to the performance of the GGA-CGT, different characteristics that interfere in the process of the crossover operator were identified and analyzed. As a result of this experimental study, knowledge of the problem domain was obtained and used to design a new high-performance crossover operator. Experimental results show that the performance of the GGA-CGT algorithm improves considerably by replacing the original crossover operator with the new one, with respect to the number of optimally solved instances and the number of generations used, with an improvement rate in effectiveness of 25% on the BPPv_u_c benchmark, considered one of the most difficult in the state-of-the-art. The GGA-CGT with the new crossover operator outperforms the performance of the best state-of-the-art algorithms and its efficiency is very high compared to the number of iterations required by those algorithms.

References

[1] Quiroz-Castellanos, M., Cruz-Reyes, L., Torres-Jimenez, J., Gómez S., C., Huacuja, H. J. F. & Alvim, A. C. (2015). A grouping genetic algorithm with controlled gene transmission for the bin packing problem. Comput. Oper. Res., 55(100), 52-64. https: //doi.org/10.1016/j.cor.2014.10.010

[2] Amador-Larrea, S., Quiroz-Castellanos, M., Hoyos-Rivera, G.-D.-J., Mezura-Montes, E., An experimental study of grouping crossover operators for the bin packing problem, Computación y Sistemas, Vol. 26, No. 2, 2022, pp. 663–682, doi:10.13053/CyS-26-2-4249

Hurricanes ocurrences and temperature forecasting by machine learning methods

Erick Estrada Patiño[a], Juan Frausto Solis[b], Guadalupe Castilla Valdez[c]

División de Estudios de Posgrado e Investigación
Tecnológico Nacional de México / Instituto Tecnológico de Ciudad Madero
Av. 1o de Mayo esq. Sor Juana Ines de la Cruz s/n Col. Los Mangos C.P.89440
[a]estrada1792@gmail.com
[b]juan.frausto@gmail.com
[c]gpe_cas@yahoo.com.mx

Climate change, among its consequences, causes alterations in natural and seasonal phenomena such as hurricanes. This phenomenon is mainly due to the accumulation of greenhouse gases, the increase in temperature on the land and sea surface, and the atmospheric alterations inherent to these changes. In Mexico, the hurricane period last approximately six months with two temporary zones: the Pacific and the Atlantic. However, in recent years, some of these phenomena have occurred outside of the temporal zone with accelerated development which causes more violent and risky phenomena for the Mexican population and its natural environment. This work aims to analyze the relationship of tropical phenomena with the temperature shown in some coastal regions of the country by analyzing climatic variables, such as land surface temperature, and applying deep neural network techniques. In other words, the objective of the work is helping to understand the relationship between these climatological phenomena and climate change.
Keywords: Climate change, Deep learning, Cyclones, Time Series Forecasting.

M3GP data refining under different scopes

Luis Muñoz Delgado [a] and Adrian Rodriguez Aguiñaga [b]

[a]Metalmecanica and [b]Sistemas Computacionales

Instituto Tecnologico de Tijuana

Calz del Tecnológico s/n, Tomas Aquino, 22414 Tijuana, B.C.

[a]lmunoz@tectijuana.edu.mx and [b]adrian.rodriguez@tectijuana.edu.mx

Abstract

In science and engineering data describe the behavior of a phenomena, for machine learning dataset establishes the features that describe the problem, which implies, the understanding of the data is to understand the problem to obtain a solution, but due to the means of data collection or due to the complexity of the problem, might generate bias or misleading dataset that are used to unwrap the problem, where the possible solutions are lost in their intrinsic relationship of datasets, since in order to understand and solve a problem some significant information is required in the dataset. In this work, M3GP is used with a function set presented in [3], which has been applied in classification [5] and regression [6] problems with success, obtaining significant results, generating new features from the original dataset, forming significant relationships that describe the problem. To uncover the refining mechanic for feature generation the following study establishes three different scopes; global, local and specific to uncover the significant impact on the behavior of dataset creation. The results show that different scopes have an impact on the complexity of new dataset and providing different scopes bring significant improvement on readability of solutions.

Table 1: Classification results for MCD3 [7]. The table values show medians of 10 runs. Global approach is using all available features to generate a solution. Local approach start by using just one feature, until all features are used. Finally Specific approach is using all features but the main objective is to reach specific fitness (Stop >= 99 Fitness), which will drive to simpler solutions (less resources needed for computation and understanding).

MCD3	Explains	Train	Test	Size	Level	Dim
Local 1	66.0	93.77	92.78	47	7	1
Local 2	29.5	99.55	95.87	173	9	6
Local 3	2.5	99.55	93.81	218	9	9
Local 4	1.1	100	90.72	246	9	9
Local 5	0.3	100	92.26	162	8	8
Global	0.2	100	92.78	163	8	8
Specific	Gen 99	99.11	94.84	37	8	6

Table 2: Classification results for Yeast [2]. The table values show medians of 10 runs. Global approach is using all available features to generate a solution. Local approach start by using just one feature, until all features are used. Finally Specific approach is using all features but the main objective is to reach specific fitness (Stop >= 94 Fitness), which will drive to simpler solutions (less resources needed for computation and understanding).

Yeast	Explains	Train	Test	Size	Level	Dim
Local 1	35.6	44.34	39.79	278	15	5
Local 2	21.0	50.49	43.69	374	16	10
Local 3	13.3	55.01	46.10	334	11	11
Local 4	9.3	60.27	49.77	224	15	11
Local 5	7.2	63.61	53.44	416	15	13
Local 6	7.0	67.35	55.50	292	14	13
Local 7	3.6	66.56	62.38	273	12	12
Global	2.8	68.33	57.11	444	12	13
Specific	Gen 94	67.05	55.27	262	13	13

Table 3: Regression results for Concrete [9]. The table values show medians of 10 runs. Global approach is using all available features to generate a solution. Local approach start by using just one feature, until all features are used. Finally Specific approach is using all features but the main objective is to reach specific fitness (Stop <= 5.9), which will drive to simpler solutions (less resources needed for computation and understanding)..

Concrete	Explains	Train	Test	Size	Level	Dim
Local 1	32.9	13.31	19.75	1564	16	29
Local 2	24.9	12.33	55.00	719	15	31
Local 3	17.8	12.00	57.69	610	13	33
Local 4	10.8	11.29	29.26	542	14	32
Local 5	10.2	9.13	24.37	543	13	30
Local 6	2.9	5.77	8.12	462	14	30
Local 7	.1	5.65	7.62	350	11	33
Global	.02	5.73	7.16	383	14	32
Specific	Gen 94	5.86	7.13	183	10	26

Table 4: Regression results for Yacht [1]. The table values show medians of 10 runs. Global approach is using all available features to generate a solution. Local approach start by using just one feature, until all features are used. Finally Specific approach is using all features but the main objective is to reach specific fitness (Stop <= .55), which will drive to simpler solutions (less resources needed for computation and understanding).

Yacht	Explains	Train	Test	Size	Level	Dim
Local 1	83.9	14.63	16.98	718	15	23
Local 2	11.6	14.65	16.90	399	11	21
Local 3	3.8	14.59	17.29	297	10	20
Local 4	.37	0.54	1.24	588	14	29
Local 5	.09	0.41	0.94	533	16	29
Global	.004	0.43	0.92	583	15	27
Specific	Gen 85	0.53	0.96	660	15	25

One of the big question of feature selection is who much information is contained in each feature, using principal component analysis the percentage of contribution of each feature can be seen on the column Explains. This way the local approach can be implemented by showing the impact in fitness quality (train and testing), readability with size, level and dimensions (Dim) of the tree. The specific approach presents a way obtain simple solution's lowering the fitness requirements to acceptable quality (remember fitness causes bloat [4], [8], this way simpler and faster solutions can be found. Finally, the last observation refers to how important each feature is, even if the explaining content is low (.004 last feature of Yacht problem) in the global scope fitness improvement can be significant to obtain the best possible solution.

References

[1] Onnink R. "Gerritsma, J. and Versluis". "geometry, resistance and stability of the delft systematic yacht hull seriest".

[2] Nakai K. Horton P. A probabilistic classification system for predicting the cellular localization sites of proteins. pages 109–115, 1996.

[3] Vijay Ingalalli, Sara Silva, Mauro Castelli, and Leonardo Vanneschi. A multi-dimensional genetic programming approach for multi-class classification problems. In Miguel Nicolau et al., editors, *17th European Conference on Genetic Programming*, volume 8599 of *LNCS*, pages 48–60, Granada, Spain, 2014. Springer.

[4] W. B. Langdon and R. Poli. Fitness causes bloat. In P. K. Chawdhry, R. Roy, and R. K. Pant, editors, *Soft Computing in Engineering Design and Manufacturing*, pages 13–22, London, 1998. Springer London.

[5] LuisMuñoz, Sara Silva, and Leonardo Trujillo. M3GP: multiclass classification with GP. In Penousal Machado et al., editors, *18th European Conference on Genetic Programming*, volume 9025 of *LNCS*, pages 78–91, Copenhagen, 2015. Springer.

[6] Luis Muñoz, Leonardo Trujillo, Sara Silva, Mauro Castelli, and Leonardo Vanneschi. Evolving multidimensional transformations for symbolic regression with m3gp. *Memetic Computing*, Sep 2018.

[7] Geological Survey. USGS geological survey.

[8] Leonardo Trujillo, Luis Muñoz, Edgar Galvan-Lopez, and Sara Silva. neat genetic programming: Controlling bloat naturally. *Information Sciences*, 333:21 – 43, 2016.

[9] I.-C. Yeh. Modeling of strength of high-performance concrete using artificial neural networks. *Cement and Concrete Research*, 28(12):1797–1808, 1998.

Feature Selection with Single Objective Particle Swarm Optimization in Gene Expression Data

José-Clemente Hernández-Hernández [a], **Gustavo-Adolfo Vargas-Hakim** [b]
David Herrera-Sánchez [a], **Efrén Mezura-Montes** [a]

[a]Artificial Intelligence Researches Institute
University of Veracruz
Campus Sur, Paseo Lote II Street, Sección Segunda, Nuevo Xalapa, 91097
Xalapa-Enríquez, Veracruz, México
{jclementehdzhdz@gmail.com, hersan19@hotmail.es, emezura@uv.mx}

[b]Systems Engineering Department
École de Technologie Supérieure
1100 Notre-Dame Street West, H3C 1K3
Montreal, Canada
{vargashakimg@gmail.com}

Abstract: Feature Selection is commonly used in Machine Learning tasks to reduce the dimensionality in databases by removing noisy and redundant features, leading to an optimal subset [1]. In this work, we introduce Feature Selection with Single Objective Particle Swarm Optimization (FS-SOPSO), a bio-inspired wrapper-based feature selection method, which uses classification accuracy as the guiding metric for the search. FS-SOPSO improves on the traditional continuous-based Particle Swarm Optimization algorithm by: (a) adding a special mutation to reduce the number of selected features, and (b) introducing a particle update to reduce the rapid convergence effect of not choosing any feature. FS-SOPSO is evaluated against traditional Feature Selection algorithms, as well as with recent bio-inspired methods [2, 3, 4], using a set of popular, high-dimensional benchmarks, resulting in consistent and significant improvements.

Keywords: Particle Swarm Optimization, Feature Selection, Machine Learning

References

[1] Xue, B., Zhang, M., Browne, W.N., Yao, X.: A Survey on Evolutionary Computation Approaches to Feature Selection. IEEE Transactions on Evolutionary Computation 20(4), 606–626 (2016). https://doi.org/10.1109/TEVC.2015.2504420

[2] Dara, S., Banka, H.: A binary PSO feature selection algorithm for gene expression data. Proceedings - 2014 IEEE International Conference on Advances in Communication and Computing Technologies, ICACACT 2014 (2015). https://doi.org/10.1109/EIC.2015.7230734

[3] Gurav, A., Nair, V., Gupta, U., Valadi, J.: Selection Using Support Vector Machines for Simultaneous Feature Selection 3, 27–37 (2015). https://doi.org/10.1007/978-3-319-20294-5

[4] Medjahed, S.A., Saadi, T.A., Benyettou, A., Ouali, M.: Kernel-based learning and feature selection analysis for cancer diagnosis. Applied Soft Computing Journal 51, 39–48 (2017). https://doi.org/10.1016/j.asoc.2016.12.010

Benchmarking GSGP-CUDA in SRBench

Jose M. Muñoz, Leonardo Trujillo, Daniel E. Hernández

Posgrado en Ciencias de la Ingeniería
Tecnológico Nacional de México/IT de Tijuana
Tijuana, BC, Mexico
mane9986@gmail.com, leonardo.trujillo@tectijuana.edu.mx, daniel.hernandezm@tectijuana.edu.mx

The first CUDA-based implementation of geometric semantic genetic programming (GSGP) was recently published by our group [1]. The algorithm, called GSGP-CUDA, is highly competitive but not equivalent with the standard sequential implementation given the different representation used. GSGP-CUDA is also at least two orders of magnitude faster in terms of computation time than standard GSGP. However, a comprehensive comparison with other state-of-the-art approaches has not yet been published. This work fills this gap, using the recently released SRBench benchmark suite [2]. SRBench provides a powerful tool that allows researchers to rigorously evaluate and compare their symbolic regression algorithms on a wide variety of synthetic and real-world problems, including for comparison several state-of-the-art techniques of both GP and non-GP methods. SRBench considers model accuracy (RMSE and R2), model complexity (size after automatic simplification) and total computation time. A limitation of SRBench, however, is that none of the GP-based approaches exploit GPU-based computation and GSGP is not considered among the tested algorithms, with our present work filling both gaps. Moreover, we extend our previous GSGP-CUDA implementation by incorporating an optimal mutation step for each individual at every mutation event.

References

[1] L. trujillo, et al., GSGP-CUDA — A CUDA framework for Geometric Semantic Genetic Programming, SoftwareX, Volume 18, June 2022, 101085.

[2] W. La Cava, et al., Contemporary Symbolic Regression Methods and their Relative Performance, Neurips Track on Datasets and Benchmarks, 2021.

A Hierarchical Approach to a Tri-objective Portfolio Optimization Problem Considering an ESG Index

Yeudiel Lara Moreno[a], Carlos Ignacio Hernández Castellanos[a]

[a]Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas
Universidad Nacional Autónoma de México
Circuito Escolar 3000, C.U., Coyoacán, 04510 Mexico City, Mexico
yeudiellm@ciencias.unam.mx, carlos.hernandez@iimas.unam.mx

The Markowitz covariance model is the classical approach to portfolio optimization. This model is a bi-objective optimization problem that aims to minimize the variance at the time that maximizes the expected return. In recent years, there has been an interest to include several other objectives to better capture the preferences of investors. Examples of these criteria are 12-month performance, 3-year performance, annual dividend, volatility, environmental, social, and corporate governance (ESG), among others [1, 2, 4].

In this work, we focus on the last one (ESG). This objective is important since more companies and individuals are starting to look at the social and environmental commitment of the companies in which to invest besides the potential profit [2, 4]. Note, that as important as ESG is, it is still considered a secondary objective since people would not consider portfolios with low expected returns or high variance even if the portfolio maximizes ESG.

Thus, we propose a hierarchical approach. First, the investor decides on allowed deterioration for the expected return and the variance (i.e. the percentual points that the investor is willing to sacrifice to look for solutions with better ESG). With this informatión, we compute the set of approximate solutions [3] considering only the classical bi-objective problem. Next, we integrate the third objective which is the ESG index. Finally, we filter any dominated solution using the three objectives.

We tested our approach with data of 40 ETFs extracted from Yahoo finance over the last year (2021-2022). The results show that the approach is capable of finding solutions that have a good ESG compromise that offers options to potential investors beyond the classical bi-objective approach. In future work, we plan to compare our formulation with the one from the state-of-the-art to better understand the advantages and disadvantages of the proposed model.

References

[1] Matthias Ehrgott, Kathrin Klamroth, and Christian Schwehm. An mcdm approach to portfolio optimization. *European Journal of Operational Research*, 155(3):752–770, 2004.

[2] Markus Hirschberger, Ralph E Steuer, Sebastian Utz, Maximilian Wimmer, and Yue Qi. Computing the nondominated surface in tri-criterion portfolio selection. *Operations Research*, 61(1):169–183, 2013.

[3] Oliver Schütze and Carlos Hernández. *Archiving Strategies for Evolutionary Multi-objective Optimization Algorithms*. Springer, 2021.

[4] Ralph E Steuer and Sebastian Utz. Non-contour efficient fronts for identifying most preferred portfolios in sustainability investing. *European Journal of Operational Research*, 2022.

A multiobjective approach for personalized inversion

Julieta Ramiro Estrada [a], Sebastián Sandoval Montiel[a], Oliver Cuate[b], Lourdes Uribe[a,b]

[a]Facultad de Ciencias Actuariales
Universidad Anáhuac México
Estado de México, México.
julieta.ramiroe@gmail.com; sebastian.sandovalmo@anahuac.mx; lourdes.uribe@anahuac.mx

[b] Escuela Superior de Física y Matemáticas
Instituto Politécnico Nacional
Ciudad de México, México.
ocuateg@ipn.mx; luriber@ipn.mx

This project stems from the need to continue promoting savings in Mexico, since according to the National Financial Education Poll, carried out by INEGI and CNBV in 2018, 21.5% of the Mexican population between 18-70 years do not have any type of savings and 31.4% only have informal savings, which means that their savings do not generate interest and are not protected by the Institute for the Protection of Bank Savings (IPAB).

Consequently, this work focuses on proposing personalized investment portfolios via clustering, highlighting two solutions when comparing traditional weight theory and evolutionary algorithms. A total of 118 students from Universidad Anáhuac México were surveyed. Then, they were categorized based on the k-modes methodology. Once the groups were created, an assets assignment based on their preference and interests was implementented. Based on the investment profiles, strategies were proposed where fixed income is included, in order to provide a portfolio that presents a fraction of the investment in Cetes and the rest in variable income, which corresponds to the assets selected in each group; this in order to differentiate each of the investments depending on their studied risk profile. For portfolio optimization, both the traditional algorithm and the multiobjective evolutionary algorithm were applied. At the same time, a sensitivity analysis was carried out in order to see the stability of certain parameters of the evolutionary algorithm. Numerical results yield that both approaches reached same indices. In addition, the Sharpe relationship was used to analyze the average return obtained in excess of the risk-free rate per unit of volatility or total risk of each of the portfolios generated by both algorithms.

In this work, a first effort was made to apply the preferences of the users to propose personalized investment portfolios using a multiobjective approach. Also, this work is a real application implemented at the university. Numerical results show both the theoretical and computational strengths of this approach. It is important to mention that this approach can be extended to a greater number of objectives, or a greater number of constraints.

References

[1] Ponsich, A., López, A., & Coello, C, A survey on Multiobjective Evolutionary Algorithms for the solution of the portfolio optimization problem and other finance and economics applications, 2013 IEEE Transactions on evolutionary computation 17,(3)

A radial hybrid estimation of distribution algorithm for the truck and trailer routing problem

Ricardo Pérez-Rodríguez[a], **Sergio Frausto-Hernández**[b]

[a]Research
CONACYT
Insurgentes Sur 1582, Crédito Constructor, Ciudad de México. C.P. 03940
dr.ricardo.perez.rodriguez@gmail.com

[b]Research
ITA
Adolfo López Mateos 1801 Oriente, Bona Gens, C.P. 20256, Aguascalientes, Ags., México
sergio.fh@aguascalientes.tecnm.mx

Abstract: The truck and trailer routing problem (TTRP) has been widely studied under different approaches. It is due to its practical characteristic that makes its research interesting. The TTRP continues being attracted to develop new evolutionary algorithms. This research details a new estimation of distribution algorithm coupled with a radial probability function from the hy-drogen. Continuous values are used in the solution representation, and every value indicates, in a hydrogen atom, the distance between the electron and the core. The key point is to exploit the radial probability distribution to construct offspring, and to tackle the drawbacks of the estimation of distribution algorithms. Various instances and numerical experiments are presented to illustrate, and to validate this novel research. Based on the performance of the proposed scheme, we can make the conclusion that incorporating radial probability distributions helps to improve the estimation of distribution algorithms.

Solving Discrete Problems with Continuous Optimization

Angel E. Rodríguez-Fernández [a]

[a]Departamento de Computación
CINVESTAV-IPN
Av. IPN No. 2508 Col. San Pedro Zacatenco, 07360, CDMX, Mexico
aerfangel@gmail.com

Many discrete optimization problems are well known in Computer Science for being NP-hard, and unless P=NP[1], we cannot design an algorithm that solves the problem exactly, in polynomial time and for any instance. One approach is to solve the problem exactly and in polynomial time, but not for any instance. The *Max-Cut* problem[1] is a well known NP-hard problem that can be solved exactly and in polynomial time considering only planar graphs.
Another way to tackle these problems is to find solutions in polynomial time and for any instance, with the property that these solutions are guaranteed to be *close* to the optimum. These are called approximation algorithms, and the idea is that for a minimization problem, the algorithm guarantees that its solution is between the optimum and a α times the optimum (with $\alpha > 1$).
Some approximation algorithms *relax* the discrete optimization problem into a continuous optimization problem (that contains the discrete version) in order to find a solution in this new continuous space. In general, this continuous solution is not feasible for the discrete problem, therefore a *rounding* is needed, i.e., this continuous solution is converted to a *close* feasible discrete solution.
This talk consists of a short summary of these types of approximation algorithms, the well known approximation algorithm for the Max-Cut problem of Goemans and Williamson[2] and a variation of it using clustering[3].

References

[1] R. M. Karp, Reducibility Among Combinatorial Problems, 1972 Miller, R.E., Thatcher, J.W., Bohlinger, J.D. (eds) Complexity of Computer Computations. The IBM Research Symposia Series. Springer. doi: 10.1007/978-1-4684-2001-2_9

[2] M. X. Goemans, et al., Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, 1995 Journal of the ACM, 42, Issue 6, pp. 1115-1145, doi: 10.1145/227683.227684

[3] A. E. Rodriguez-Fernandez, et al., Clustering Improves the Goemans–Williamson Approximation for the Max-Cut Problem, 2020 Computation, 8, pp. 75, doi: 10.3390/computation8030075

A Generalized Linear Model estimation for agricultural variables by an extension of the so-called Recursive Least Square Method

R. Juárez Del Toro[a], **V. Azhmyakov**[b], **C. Borrego Salcido**[c], **A. S. Fonseca Zendejas**[d]

[a]Instituto Tecnológico Superior de Lerdo
Tecnológico Nacional de México Campus Lerdo
Ciudad Lerdo, Durango. México.
raymundo.jt@itslerdo.edu.mx

[b]School of Computer Engeneering
National Research University
Moscow, Russia.
v.azhmyakov@hse.ru

[c]Unidad Interdeisciplinaria de Ingeniería campus Coahuila
Instituto Politécnico Nacional
San Buenaventura, Coahuila. México.
cborregos@ipn.mx

[b]Departamento de Economía
Universidad de Loyola
Sevilla, España.
asfonseca@uloyola.es

Abstract

The Linear Square LS and Recursive Linear Square RLS algorithms are implemented in order to identify the linear model under uncertainties for crop yield, weather and plant biometry variables in a local farm in Aguascalientes, México. The conventional broccoli represents 39% of the whole farm production. The model identification allows to forecast the Crop Yield information before harvesting, and to plan, implement, manage and control the harvesting strategy and its resources: human or material.

Keywords: Recursive Least Square, Crop Yield, Generalized Linear Models, Estimation, Weather, Plant Biometry.

A Continuation Method for the Treatment of Degenerated Multi-objective Optimization Problems

Oliver Cuate[a], Oliver Schütze[b]

[a]ESFM
IPN
ocuateg@ipn.mx

[b]Computer Science Department
Cinvestav-IPN
schuetze@cs.cinvestav.mx

Multi-objective continuation methods are very powerful tools for the numerical treatment of continuous multi-objective optimization problems (MOPs). All of these methods, however, are based on certain regularity assumptions on the model that imply that the solution set of the given MOP – the Pareto set, respectively its image, the Pareto front – is not degenerated leading to an exessive computational cost or even failure when applied to problems which are at least locally degenerated.
In this talk, we present and discuss a new predictor step for the use within multi-objective continuation methods that automatically detects (numerical) degeneration, and that can perform movements in "essential" directions at a given candidate solution. We will further integrate the this new prector in the recently proposed Pareto Tracer, and will show on selected benchmark problems that this new continuation method can efficiently handle both degenerated and non-degenerated MOPs.

Combination of WRF numerical model and ARIMA model to improve wind speed forecasts

Bernabé-Morales M.[a], Cadenas-Calderón E.[a], Campos-Amezcua R.[b] Garcia-Rodriguez I.[a]

[a]Faculty of Mechanical Engineering
Universidad Michoacana de San Nicolás de Hidalgo
Santiago Tapia 403, Centro, 58000 Morelia, Mich.
091558d@umich.mx, ecadenas@umich.mx, 9602418a@umich.mx

[b]National Center for Technological Research and Development
Tecnológico Nacional de México
Interior Internado Palmira S/N, Col. Palmira, Cuernavaca 62490, México
rafael.ca@cenidet.tecnm.mx

Wind as a source of electricity has several characteristics that differentiate it from traditional energy sources. The wind is not naturally dispatchable due to its intermittence and variable nature. An accurate wind forecast can help to reduce the level of uncertainties in wind production and allow system operators to plan ahead for the varying wind outputs [1]. This paper presents the conformation of a model destined to forecast the wind speed 24 hours ahead. The model was built as a combination of two techniques, the numerical WRF technique that was used for the 24-hour wind speed forecast in Michoacan, Mexico, and the statistical ARIMA technique training with a time series from Tepuxtepec, Michoacan, Mexico. The statistical technique was integrated into the numerical technique to improve the forecast accuracy. The forecasted results show that the combined WRF-ARIMA technique outperforms individual techniques by up to 18%, demonstrating the positivity of combining both techniques.

Keywords: Wind forecast, ARIMA, WRF.

References

[1] Xiaoguang L., Estimating Wind Forecast Errors and Quantifying Its Impact on System Operations Subject to Optimal Dispatch, University of Toronto, 2011, `https://tspace.library.utoronto.ca/bitstream/1807/31312/3/Li_Xiaoguang_201111_MASc_thesis.pdf`

Enhanced combinatorial forecasting for Mexican Stock Exchange

José L. Purata-Aldaz[a] , **Juan Frausto-Solis**[a]
Guadalupe Castilla-Valdez[a], **Juan Javier González-Barbosa**[a]

[a]División de Estudios de Postgrado e Investigación
Tecnológico Nacional de México/IT Ciudad Madero
Cd. Madero, Tamaulipas, México
jpurata.itcm@gmail.com

Forecasting is an ancient practice used to predict temperature, flowing, planting, harvesting, production, and finance. Financial forecasting held the attention of practitioners, investors, and analysts for years. The forecast's main objective is to provide the lowest forecast error and the best accuracy, for this, there are a variety of tools that allow modeling forecasts like Auto-Regressive, Moving-Average, Long Short-Term Memory, Neural Networks, and Exponential Smoothing. In 1969, Bates and Granger proposes to combine forecasts to reduce errors instead to discard one less favorable. This work aims to analyze the combinational forecasting method and its optimization with a heuristic algorithm that improves results in small horizons using assets from Mexican Stock Exchange (Bolsa Mexicana de Valores, BMV). Results show that enhanced combinatorial forecasting shows better results than simple forecasting and the equally weighted combination.

Keywords: Time series forecasting, heuristics, optimization, financial.

1 Introduction

Forecasting is a complex duty, and all methods try to reduce errors and increase accuracy to make better decisions. After using several forecast models, analysts and businessmen choose only one, the best, and discard the rest. Barnad in 1963, shows a combination of methods to develop better quality controls [2], Bates & Granger propose not discarding bad methods, instead combining methods to obtain better results [3], [4].

In case it is required to combine 2 or 3 forecasting methods, it is relatively easy to weigh the methods, as average of each one [9], however, the situation is complicated when there are 4 or more methods, as well as a large number of time series to model. Therefore, it is necessary to apply tools that allow weighting between several methods that allow minimizing the forecast error.

Threshold Accepting (TA) is an algorithm proposed by Duek and Scheuer [5], this algorithm can accept poor or suboptimal solutions under the acceptance criterion. In the beginning, at high temperatures, poor solutions can be easily accepted and at low temperatures, fewer unsatisfactory solutions can be accepted. TA is a tool that has proved efficient in applications like [8], [7] and [10].

Figure 1: Classical Threshold Accepting algorithm

Initialize: n_{steps}, α, convergence = false; k= 1
  Compute threshold sequence T_r. Compute $f(x^{old}) = f(x^r)$.
    Randomly generate a current solution $x^{new} \varepsilon X$
      while convergence = false **do**
        for $i = 1: n_{steps}$ **do**
          Generate $x^n \in N(x^c)$ and compute $\Delta = f(x^{new}) - f(x^{old})$
            if $\Delta < T_k$ **then** $x^{new} = x^{old}$
        end for
          $k = k + 1$
          $T_k = \alpha T_k$
          if $\Delta \leq e$, or $k \geq N_{max}$, convergence = true
        end while
$x^{sol} = x^{old}$

This work formulates an application of a heuristic algorithm: Threshold Accepting (TA), as a solution to establish a weighing of each method in a combinatorial strategy to use 6 forecast methods and 47 stocks in BMV.

2 Proposed method

Several forecasting methods are combined to obtain the minimum error that is better than the individual error of each method [1]. The forecast methods used in this work are: Exponential Smoothing STate space (ETS), Regressive Integrated Moving Average (ARIMA; Auto-ARIMA to determine the best ARIMA method), Seasonal Decomposition of Time Series by Loess (STL), Neural Network forecasting method (NNetar), Exponential Smoothing State Space-Box-Cox-ARMA (TBATS) and the AutoRegressive Fractionally Integrated Moving Average (ARFIMA). For all these methods are trained and tested to get the error of each one.

$$\begin{aligned} &\min err = \sum_{i=1}^{M} w_i \hat{Y}_i \\ &\text{Subject to } \textstyle\sum_{m=1}^{M} w_i \leq 1; \forall w_i \geq 0; i = 1 \ldots M \end{aligned} \tag{1}$$

Where err: estimated error for combined forecasting methods, w_i: weight for each method, $\hat{Y}_i = f(M_i, \hat{Y}_{t+h,i})$: predict h periods from M_i forecast method.
The error metric used is the symmetric mean absolute percentage error (sMAPE), a variation of the classic mean absolute percentage error (MAPE) metric [6].

$$sMAPE = \left[\frac{2}{h} * \sum_{t=n+1}^{n+h} \frac{|Y_i - \hat{Y}_i|}{|Y_i| + |\hat{Y}_i|}\right] * 100\ (\%) \tag{2}$$

Where Y_i is the current observation of the time series; the variable $\hat{Y}_i$ represents the predicted value, n is the amount of data in the time series, and h represents the length of forecast horizon.

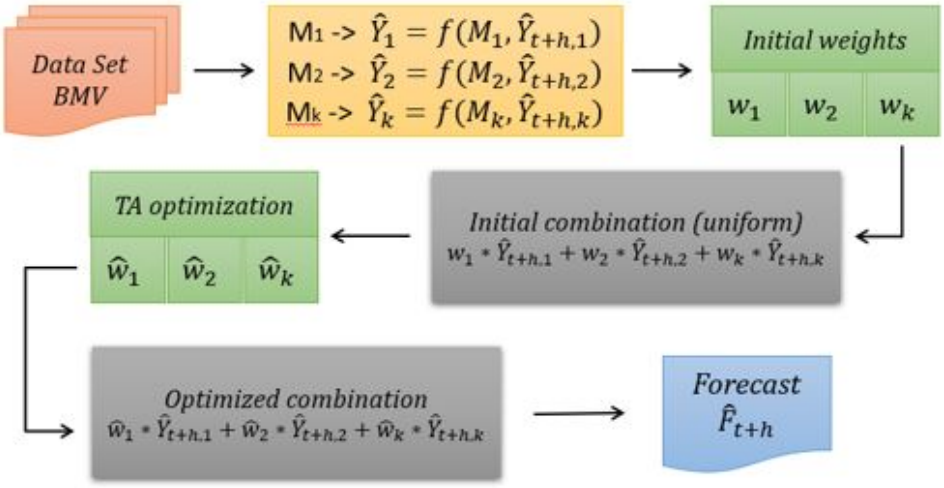

Figure 2: Methodology for forecast combination

3 Results

Used datasets consist of daily time series from BMV. For the forecast, we use daily close price for 47 assets from 01-01-2020 to 01-08-2022 (667 obs.), setting 90 percent of these for training and 10 percent to test set. Each asset is analyzed by each method getting sMAPE after a uniform combination is formulated (equally weight) applying (1) and it is stored as an initial solution for the TA algorithm. In the optimization process (2) is used as an objective function and the result of this process is the weight of each method for the optimal combination. Results show that after 30 repetitions of the described process, the optimized combination, on average, shows a smaller sMAPE in comparison to individual methods.

Table 1: Average sMAPE of each forecasting method, uniform combination, and optimized combination

ETS	ARIMA	STL	NNAR	TBATS	ARFIMA	Cmb Un	Cmb Opt
2.605	2.607	7.341	2.593	2.587	3.591	3.272	1.955

4 Conclusions and future works

This work demonstrates that a selection of good accuracy methods and his combination results in a very good performance in forecasting for stock price analysis. In future work, we see the opportunity to implement other forecasting methods, tools to improve TA tuning and techniques to avoid stagnation and the use of datasets from SP500 and other relevant stock markets.

Acknowledgments: The authors thanks to Consejo Nacional de Ciencia y Tecnologia (CONACYT) for the scholarship and Tecnológico Nacional de México for graduate studies.
Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Armstrong, J. S. (2001). *Principles of forecasting: a handbook for researchers and practitioners*, Volume 30. Springer.

[2] Barnard, G. A. (1963). New methods of quality control. *Journal of the Royal Statistical Society. Series A (General) 126*(2), 255–258.

[3] Bates, J. M. and C. W. Granger (1969). The combination of forecasts. *Journal of the operational research society 20*(4), 451–468.

[4] Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International journal of forecasting 5*(4), 559–583.

[5] Dueck, G. and T. Scheuer (1990). Threshold accepting: A general purpose optimization algorithm appearing superior to simulated annealing. *Journal of computational physics 90*(1), 161–175.

[6] Flores, B. E. (1986). A pragmatic view of accuracy measurement in forecasting. *Omega 14*(2), 93–98.

[7] Frausto-Solis, J., L. Hernández-Ramírez, G. Castilla-Valdez, J. J. González-Barbosa, and J. P. Sánchez-Hernández (2021). Chaotic multi-objective simulated annealing and threshold accepting for job shop scheduling problem. *Mathematical and Computational Applications 26*(1), 8.

[8] Frausto Solis, J., J. L. Purata Aldaz, M. González del Angel, J. González Barbosa, and G. Castilla Valdez (2022). Saipo-taipo and genetic algorithms for investment portfolios. *Axioms 11*(2), 42.

[9] Hyndman, R. J. and G. Athanasopoulos (2018). *Forecasting: principles and practice*. OTexts.

[10] Winker, P. and D. Maringer (2007). The threshold accepting optimisation algorithm in economics and statistics. In *Optimisation, econometric and financial analysis*, pp. 107–125. Springer.

A bibliometric analysis of texts on homophobia, transphobia, HIV discrimination, and gender violence from 2012 to 2022.

Luis Ángel Meza-Zárate[a], Roman Anselmo Mora Gutiérrez[b], Edwin Montes-Orozco[c], Juan Villegas-Cortez[b], Jorge Mercado-Mondragon, Jorge Juárez-Gómez, Bibiana Obregón Quintana.

[a] División de ciencias y Artes para el Diseño. UAM-Azcapotzalco [b] Departamento de Sistemas. UAM-Azcapotzalco [c] Departamento de Matemáticas Aplicadas y Sistemas. UAM-Cuajimalpa [d] Departamento de Sociología. UAM-Azcapotzalco [e] Departamento de Quimica. UAM-Iztapalapa [f] Facultad de Ciencias. UNAM

Abstract

In this work, we present a bibliometric analysis using the complex networks approach by solving a bi-objective variant of the optimization problem known as semi-coloration (SCOP) to identify currents of thought (communities) in the literature on homophobia, transphobia , HIV discrimination and gender violence.
The numerical results show that four strong communities are formed in the modeled networks.A bibliometric analysis is presented in this paper. A complex network model and a variant bi-objective of the semi-coloring optimization problem (SCOP) are used to identify currents of thought (communities) in the literature about homophobia, transphobia, HIV discrimination, and gender violence topics. We can say that four firmly communities are identifiable based on numerical results.

1 Introduction

Currently, the human being is immersed in a hyperconnected world in which the generation of information on any subject increases exponentially. For this reason, on many occasions, researchers must construct and consult bibliographic and bibliometric analyzes on the subject of their interest in order to identify currents of thought, ideas and relevant authors.
In [1] is mentioned that the analysis of scientific publications constitutes an essential link in the research process; since it can qualify the knowledge-generating process quality and the impact.
In [2, 3, 4], a complex network model has been used for bibliometric analysis. In this work, the identification and determination of communities are used to determine currents of thought.

2 Methodology

Initially, 15 papers about each topic analyzed (homophobia, transphobia, HIV discrimination, and gender violence) were collected from the ScienceDirect repository. Each paper selected was generated from 2012 to 2022; it is open access.
Afterward, each paper was analyzed by a word algorithm to determine 20 words more frequently. Next, we built a complex network model in which the authors and ideas related child. The visibility method is employed to relate two ideas [5, 6].
Then, we use the variant of SCOP, which is shown in equation 1 . In general terms, this variant of SCOP can be defined as given an undirected network $R : (V, E)$ with weights on the edges; given the

network $V : C_1 \cup C_2 \cup C_k \cup S$ (C_i is formed by the nodes colored by the ith value, while the uncolored nodes form S) is found, such that the sum of the edges of the vertices with the same color in the original network (h) is maximized, and the sum of the edges of the uncolored vertices g is minimized

$$\begin{array}{rcl} \max(z_1) & : & h \\ \min(z_2) & : & g \\ \min(\sum w_{i,j}e_{i,j,1}, \sum w_{i,j}e_{i,j,2}, \ldots \sum w_{i,j}e_{i,j,k}) & = & h \\ \sum w_{i,j}e_{i,j,s} * |S| & = & g \\ v_{i,k}\bar{e}_{i,j,k} + v_{j,k}\bar{e}_{i,j,k} & \leq & 1 \\ C_1 \cup C_2 \cup C_k \cup S & = & V \\ C_i \cap C_j & = & \emptyset \\ C_i \cap S & = & \emptyset \\ \sum v_{i,k} & \geq & 1 \\ v_{i,k} \in [0,1] \text{ and it is integer } \ \forall i = 1,2,...,n \ \forall k = 1,...,K & & \\ \text{each subgraphs } C_k \text{ is connected} \forall k = 1,...,K & & \end{array} \tag{1}$$

3 Results

Based on communities identified the structure of complex networks, we can say there are four currents of thought in the literature consulted.

References

[1] Alfaro, A. C, Sánchez, M. D. L. T, Aragón, H. A. P, & Martínez, M. P. (2020). Análisis bibliométrico sobre la responsabilidad social social universitaria publicados en revistas que pertenecen a Scopus durante el periodo 2016-2018. Aglala, 11(S-1), 257-283.

[2] Tian, X., Geng, Y., Zhong, S., Wilson, J., Gao, C., Chen, W., ... & Hao, H. (2018). A bibliometric analysis on trends and characters of carbon emissions from transport sector. Transportation Research Part D: Transport and Environment, 59, 1-10.

[3] Lozano, S., Calzada-Infante, L., Adenso-Díaz, B., & García, S. (2019). Complex network analysis of keywords co-occurrence in the recent efficiency analysis literature. Scientometrics, 120(2), 609-629.

[4] Zhang, Y., & Zhou, Y. (2017, July). Bibliometrics analysis of complex networks research. In Proceedings of the International Conference on Business and Information Management (pp. 24-28).

[5] Lacasa, L.; Luque, B.; Ballesteros, F.; Luque, J.; Nuño, J.C., 2008. From time series to complex networks: The visibility graph. Proc. Natl. Acad. Sci. USA, 105, 4972–4975. DOI:10.1073/pnas.0709247105

[6] Guzmán Vargas L,, Aguilar Velázquez, D., Obregón Quintana, B., (2016) ANÁLISIS DE ORGANIZACIÓN Y CORRELACIONES EN TEXTOS ESCRITOS Boletín UPIITA Numero 54 IPN. México

On the Effect of Temporal Heterogeneity on Selection Pressure of evolutionary algorithms

Victor Manuel Sanchez Sanchez, Carlos Gershenson Garcia,
Carlos Ignacio Hernández Castellanos

Universidad Nacional Autonoma de Mexico
Instituto de Investigaciones en Matemáticas Aplicadas y Sistemas
Escolar 3000, C.U., Coyoacán, 04510 Ciudad de México, CDMX
vicool2@comunidad.unam.mx
cgg@unam.mx
carlos.hernandez@iimas.unam.mx

Many of the evolutionary techniques apply highly elitist methodologies, which breaks the balance between exploration and exploitation. In many cases, this leads to a loss of diversity and missing promising solutions. Thus, it is interesting to study evolutionary algorithms that apply adaptive parameters to improve the balance between exploration and exploitation. In particular, we are interested on studying the effect of temporal heterogeneity on the selection pressure.

Temporal heterogeneity is a property of a system where the most important individuals in the population change more slowly than the less important ones (which allows us to generate diversity) and to find solutions in less time by trying to minimize the number of computational instructions.

This concept was used in [1] in the context of evolutionary algorithms. The authors used it to modify rates of crossover and mutation according to their fitness. Their results showed that temporal heterogeneity improved the results on the academic problems selected.

In this work, we aim to study the impact of temporal heterogeneity on the search and to find good values to balance exploration and exploitation. We have analyzed genetic algorithms, differential evolution, and evolutionary strategies on five classical academic problems. Our preliminary results indicate that there is a significant statistical difference in more than 80% of the cases, measured by the Kruskal-Wallis test and Adhoc postprocessing. This results indicates that we should adapt the selection pressure at different stages of the search. As future work, we intend to utilize this information to develop new adaptive rules to set the selection pressure during a run of the algorithm as well as extend our study to multi-objective optimization.

References

[1] Martínez Arévalo Yoshio Ismael, La heterogeneidad temporal mejora los algoritmos de búsqueda, Universidad Nacional Autónoma de México 2021

A Bounded Archiver for Hausdorff Approximations of the Pareto Front for Multi-Objective Evolutionary Algorithms

Carlos Hernández[a], **Oliver Schütze**[b]

[a]IIMAS
UNAM
carlos.hernandez@ciencias.unam.mx

[b]Computer Science Department
Cinvestav-IPN
schuetze@cs.cinvestav.mx

Multi-objective evolutionary algorithms (MOEAs) have been successfully applied for the numerical treatment of multi-objective optimization problems (MOP) during the last three decades. One important task within MOEAs is the archiving (or selection) of the computed candidate solutions, since one can expect that an MOP has infinitely many solutions. We present and analyze in this talk ArchiveUpdateHD, which is a bounded archiver that aims for Hausdorff approximations of the Pareto front. We show that the sequence of archives generated by ArchiveUpdateHD yields under certain (mild) assumptions with a probability of one after finitely many steps a Δ_+-approximation of the Pareto front, where the value Δ_+ is computed by the archiver within the run of the algorithm without any prior knowledge of the Pareto front. The knowledge of this value is of great importance for the decision maker, since it is a measure for the "completeness" of the Pareto front approximation. Numerical results on several well-known academic test problems as well as the usage of ArchiveUpdateHD as an external archiver within three state-of-the-art MOEAs indicate the benefit of the novel strategy.

On the Study of Some Design Aspects for Evolutionary Algorithms for the Generation of Boolean Functions

Sara Mandujano[a], Dra. Adriana Lara[a], Dr. Juan Carlos Ku Cauich[b]

[a]Escuela Superior de Física y Matemáticas del
Instituto Politéctino Nacional
Av. Instituto Politéctino Nacional s/n,
Edificio 9 Unidad Profesional "Adolfo López Mateos"
smandujanov2000@alumno.ipn.mx, alaral@ipn.mx

[b]Computer Science
CINVESTAV-IPN
Av. IPN 2508, San Pedro Zacatenco
Mexico City 07300, Mexico
jcku@cs.cinvestav.mx

Abstract. Boolean functions have become a key element in modern cryptography, thus triggering a need to develop methods that efficiently construct Boolean functions with good cryptographic properties. Two desired characteristics of Boolean functions are balancedness and high non-linearity. Combining these features makes these functions more resistant to cryptanalytic attacks. More recently, there has been an increased interest in Weight-wise Perfectly Balanced (WPB) Boolean functions used in the FLIP family of stream ciphers [2][1]. One approach that has been considered to develop such functions is Genetic Algorithms (GAs) [1][3]. This work investigates the complexity and influence of design aspects when using GAs to develop WPB Boolean Functions, e.g., population setting, constraint management, and local search. The aim is to identify the most complex stages for this particular problem to refine them for a smoother application of the technique and thus tackle larger problem sizes.

Keywords: Boolean Functions · genetic algorithms · cryptography.

References

[1] Luca Mariot, Stjepan Picek, Domagoj Jakobovic, Marko Djurasevic, and Alberto Leporati. Evolutionary construction of perfectly balanced boolean functions. *arXiv preprint arXiv:2202.08221*, 2022.

[2] Pierrick Méaux, Anthony Journault, François-Xavier Standaert, and Claude Carlet. Towards stream ciphers for efficient fhe with low-noise ciphertexts. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 311–343. Springer, 2016.

[3] Stjepan Picek, Domagoj Jakobovic, Julian F Miller, Elena Marchiori, and Lejla Batina. Evolutionary methods for the construction of cryptographic boolean functions. In *European Conference on Genetic Programming*, pages 192–204. Springer, 2015.

An experimental study of the Grouping Genetic Algorithm for Cooperative Co-evolution Variable Decomposition in Large-scale Optimization Problems

Guadalupe Carmona-Arroyo, Marcela Quiroz-Castellanos and Efrén Mezura-Montes

Artificial Intelligence Research Institute
Universidad Veracruzana
Xalapa, Veracruz, Mexico
gcarmonaarroyo@gmail.com, maquiroz@uv.mx, emezura@uv.mx

When optimization problems have a high number of variables the algorithms that solve them present difficulties to do it. These types of problems are known in the literature as Large-scale Optimization Problems. To deal with these problems there is a well-known approach in the state-of-the-art which is called Cooperative Co-evolution, this method uses the idea of divide and conquer, creating subsets or subgroups of variables to optimize separately and then create the solution to the complete problem.
The creation of these aforementioned subgroups, also called variable decomposition, is another problem to solve since the variables that interact with each other must remain together in the same subgroup, just as the variables that do not present interaction can be in separate groups. Therefore, the variable decomposition can then be seen as a combinatorial optimization problem.
One of the best strategies to optimize groups creation is the use of Grouping Genetic Algorithms (GGAs), which are Genetic Algorithms that work under a group-based representation.
Previously, we created a Grouping Genetic Algorithm to obtain an adequate variable decomposition of the problem, this algorithm was tested in 18 functions of a constrained benchmark. After that, by analyzing the results of our algorithm, we determined that we could improve the results by choosing better components of the algorithm, such as selection, replacement, reproduction operators, and so on.
In this work, we present an analysis of the performance of the GGA creating alternative versions of it, in which we individually study the contribution of its components to the algorithm performance.
The results show that the algorithm can be improved, for example, by taking strategies that encourage the replacement of the worst individuals after the crossover, or by performing a mutation of the individuals in which the repair of the solutions is allowed through the integration of variables in previously created groups. Also, this new version of the GGA was evaluated through the Cooperative Co-evolution approach showing competent results.
The experiments carried out also indicate that the parameters such as crossover percentage and mutation percentage can be adjusted to a better performance of the Grouping Genetic Algorithm.

References

[1] Potter, M. A., & De Jong, K. A. (1994, October). A cooperative coevolutionary approach to function optimization. In *International Conference on Parallel Problem Solving from Nature* (pp. 249-257). Springer, Berlin, Heidelberg.

[2] Carmona-Arroyo, G., Quiroz-Castellanos, M., & Mezura-Montes, E. (2022). Variable Decomposition for Large-Scale Constrained Optimization Problems Using a Grouping Genetic Algorithm. Mathematical and Computational Applications, 27(2), 23.

A Grouping Genetic Algorithm for RGB-image segmentation

Norma-Angélica Zavaleta-García [a], Oscar Fernández-Solano [a], Marcela Quiroz-Castellanos [a], Héctor-Gabriel Acosta-Mesa [a], Efrén Mezura-Montes [a]

[a] Artificial Intelligence Research Institute Universidad Veracruzana Campus Sur, Calle Paseo Lote II, Sección 2a, N° 112, Nuevo Xalapa, Xalapa, Ver., México 91097

zS21000452@estudiantes.uv.mx, zS21000458@estudiantes.uv.mx, maquiroz@uv.mx, heacosta@uv.mx, emezura@uv.mx

The problem of RGB-image segmentation for a fixed number of segments can be considered as a type of clustering problem, where we want to partition an image into a fixed number of parts or sets under an arrangement heuristic according to the similarity among the objects and all those belonging to a cluster [2]. In machine learning, clustering is considered one of the most challenging problems. From the optimization point of view, clustering problems are NP-hard, which motivates the search for algorithms and the use of meta-heuristics, particularly the use of evolutionary algorithms, which are effective for solving these kinds of problems [4]. At the same time, the image segmentation problem can be seen as a grouping problem where the pixels are taken as the items and the segments of the image are the groups. According to this, it is natural to think about the use of Grouping Genetic Algorithms (GGAs) to solve the aforementioned problem. This type of algorithms has been used to solve clustering problems achieving competitive performances [5].
Some GGAs have been proposed for the clustering problem in [3], [6], [1], but they didn't address the RGB-image segmentation problem. Therefore, in this work a GGA is proposed to solve the image segmentation problem in the RGB color space, assuming that the number of segments is known. The proposed GGA was evaluated by comparing the best solutions found against those obtained using the K-means algorithm, where the difference between them is close to zero. This suggests that the proposed algorithm has a competitive performance for RGB-image segmentation.

References

[1] Agustı, L., Salcedo-sanz, S., Jiménez-fernández, S., Carro-calvo, L., Delser, J. & Portilla-figueras, J. A new grouping genetic algorithm for clustering problems. *Expert Systems With Applications*. **39**, 9695–9703 (2012)

[2] Halder, A., Pramanik, S. & Kar, A. Dynamic image segmentation using fuzzy c-means based genetic algorithm. *International Journal Of Computer Applications*. **28**, 15–20 (2011)

[3] Hong, T., Lin, F. & Chen, C. Using the group genetic algorithm for attribute clustering. *2012 IEEE Congress on Evolutionary Computation*, 1–5 (2012).

[4] Hruschka, E., Campello, R. & Freitas, A. A survey of evolutionary algorithms for clustering. *Ieee Transactions On Systems, Man, And Cybernetics, Part C (applications And Reviews)*. **39**, 133–155 (2009)

[5] Ramos-figueroa, O., Quiroz-castellanos, M., Mezura-montes, E. & Schütze, O. Metaheuristics to solve grouping problems: A review and a case study. *Swarm And Evolutionary Computation*. **53** pp. 100643 (2020)

[6] Yusoh, Z. & Tang, M. Clustering composite SaaS components in cloud computing using a grouping genetic algorithm. *2012 IEEE Congress on Evolutionary Computation*, 1–8 (2012)

Learning explainable classification models by approximating the minimum clique partition problem

Saul Neri-Ortiz[a], Miriam Pescador-Rojas[b], Salvador Godoy-Calderón[c]

Centro de Investigación en Computación
Instituto Politécnico Nacional
Ciudad de México, México
[a] snerio2021@cic.ipn.mx
[c] sgodoyc@cic.ipn.mx
Escuela Superior de Cómputo
Instituto Politécnico Nacional
Ciudad de México, México
[b]mpescadorr@ipn.mx

In graph theory, a clique C of a given graph G is a fully connected subgraph. The problems of finding a maximal clique, that is, a clique that is not included in a larger clique, and finding a minimum clique partition are problems belonging to the NP-Hard complexity class. However, said problems have multiple interesting applications in computer science and other disciplines, such as bioinformatics, communication systems, and social networks. We propose a machine learning algorithm for classification that models a given dataset into a special kind of rejectability graph and produces a set of classification rules expressed using Boolean logic. The algorithm works by iteratively finding a clique from the rejectability graph and extracting a Boolean function out of it, until it creates a complete clique partition of the initial graph. To improve the explainability of the model, finding the fewest number of Boolean clauses is desired. As a consequence, the algorithm aims to find the minimum clique partition by repeatedly finding a maximal clique and removing it from the rejectability graph. Given that the time complexity for a brute force approach to solving those problems grows exponentially, using classical algorithms is impractical. To navigate the large search space for the maximal clique problem, our proposal uses a modification to the traditional Artificial Ant Colony (ACO) and a Genetic Algorithm for finding good approximations using reasonable resources within a practical time range.

Machine learning techniques to predict the life testing of the fixed offshore platform structure

Youness El Hamzaoui[*], *Juan Antonio Álvarez Arellano*
and
Rodrigo Daniel Álvarez Bello Martínez

Facultad de Ingeniería, Universidad Autónoma del Carmen, Campus III, Avenida Central S/N, Esq. con Fracc. Mundo Maya, C.P. 24115, Ciudad del Carmen, Campeche, México.

** Corresponding author: eyouness@pampano.unacar.mx;*

Abstract

The fixed offshore platforms in Mexico produce about 1.90 million barrels of oil daily, with a value of 1.5 million dollars, placing Mexico as an essential oil country worldwide. Nevertheless, there is a lack of a general analytic method to calculate the useful life of structures due to ensuring its structural integrity is very complex because of the factors it is subjected to, such as operational, meteorological, extreme storm, seismic and tabular joint fatigue loads. Since the process depends on several factors, the modeling of these processes requires many problems showing nonlinear behavior, which are difficult to describe by a linear mathematical model taking into account the multivariate system. This process couldn't be solved by simple linear multivariate correlation. However, the developments of machine learning originated from artificial intelligence and are now used in various fields in reliability engineering and life testing, due to their powerful nonlinear modeling

capability. This study presents machine learning techniques that can be used to predict the life testing of fixed offshore platform structures. We used a real-world dataset to train and test several machine learning models, including support vector machines, decision trees, neural networks, and random forests. We compared the accuracy of the predictions made by each model and found that the random forest model performs the best. The proposed algorithm has been validated by comparing experimental measurement and simulated analysis. Many statistical parameters such as determination coefficient (R^2), adjusted R^2, p-value, F value, coefficient of variation (C.V%), adequate precision, and statistically significant results were presented showing strong agreement with a very low percentage of error (0.25%), high Pearson correlation (R^2=0.9979), and short computing time (less than five minutes).

The results for the proposed methodology were efficient and attractive in a reasonable way to give a satisfactory picture of the structural behavior during the different load stresses and evaluate the present settlements to conclude its structural integrity. Therefore, it could be concluded that machine learning can be used effectively to predict and provide more accurate predictions of the fatigue life of a platform structure.

Keywords: Machine learning; Random forest model; Training approach; Life testing, Fixed offshore platform.

Estimating Evapotranspiration of Mexican Tropical Climate using Machine Learning Algorithms

Youness El Hamzaoui[1], Jose Luis Vazquez Avila[1], Juan Israel Yañez Vargas[2], Juan Antonio Álvarez Arellano[1], Carlos Roman De la Cruz Dorantes[3]*

[1]Facultad de Ingeniería, Universidad Autónoma del Carmen, Campus III, Avenida Central S/N, Esq. con Fracc. Mundo Maya, C.P. 24115, Ciudad del Carmen, Campeche, México.
[2]Universidad Politécnica de Juventino Rosas, Departamento de Ingeniería en Redes y Telecomunicaciones. Miguel Hidalgo 102, Comunidad de Valencia, 38253 Juventino Rosas, Gto.
[3]Facultad de Ciencias de la Información, Universidad Autónoma del Carmen, Campus "José Ortíz Ávila" Calle 56 por Av. Concordia, CP 24180. Ciudad del Carmen, Mexico

** Corresponding author: eyouness@pampano.unacar.mx;*

Abstract

The measurement of evapotranspiration (ET) is the most important factor in irrigation scheduling. Evapotranspiration is the loss of water from the surface of plant and soil. The amount of (ET) varies with the climate. There is no one-size-fits-all method for calculating the Evapotranspiration, because there are many factors that can affect its longevity. These include things like inside & outside temperature, humidity, barometric pressure, wind speed & direction, dew point and rainfall. Because of this, it is difficult to create a mathematical model that accurately describes all the possible outcomes. Since the process depends on several factors, the modeling of these processes requires many problems showing nonlinear

behavior, which are difficult to describe by a linear mathematical model taking into account the multivariate system. This process couldn't be solved by simple linear multivariate correlation. However, the developments of machine learning originated from artificial intelligence and are now used in various fields in in the environment, water cycle, climate, atmosphere and evapotranspiration, due to their powerful nonlinear modeling capability.

This study aimed to estimate the evapotranspiration of Mexican tropical climate using machine learning algorithms. The dataset used was acquired from the Davis Vantage Pro2 as a full weather station and the National Water Commission (CONAGUA) to training and testing several machine learning models, including support vector machines, decision trees, neural networks, and random forests. We compared the accuracy of the predictions made by each model and found that the random forest model performs the best. The proposed algorithm has been validated by comparing experimental measurement and simulated analysis. Many statistical parameters such as determination coefficient (R^2), adjusted R^2, p-value, F-value, coefficient of variation (C.V%), adequate precision, and statistically significant results were presented showing strong agreement with a very low percentage of error (0.35%), high Pearson correlation (R^2=0.9825), and short computing time (less than three minutes).

The study found that the Random Forest algorithm was more accurate in estimating evapotranspiration in Mexican tropical climate, with a mean absolute error of 0.085 mm/day. This algorithm can be used in future studies to more accurately estimate evapotranspiration in other tropical climates.

Keywords: Machine learning; Random forest algorithm; Training approach; Validation approach; Statistical Metrics; Evapotranspiration.

Emotion face recognition utilizing evolutionary CNN

Edwin Bryan Salas López [a], Juan Villegas Cortez [b], Silvia Beatriz González Brambila [c], Francisco Fernandez de Vega [d], Graciela Roman Alonso [e], Salomón Cordero Sánchez [f]

[a] [b] [c]Departamento de Sistemas
Univesidad Autónoma Metropolitana, Unidad Azcapotzalco
Av. San Pablo 180, Col. Reynosa, C.P. 02200 Mexico City, Mexico
{al2213800574, juanvc, sgb}@azc.uam.mx

[d] Department of Computer Science
University of Extremadura
Merida, Spain
fcofdez@unex.es

[e]Departamento de Ingeniería Eléctrica
Universidad Autónoma Metropolitana, Unidad Iztapalapa
Mexico City, Mexico
grac@xanum.uam.mx

[f]Departamento de Química
Universidad Autónoma Metropolitana, Unidad Iztapalapa
Mexico City, Mexico
scs@xanum.uam.mx

Advances in machine learning and its relationship with data and information have made it possible for computers to use facial expressions for emotion analysis and recognition, although it lacks the ability to extract full and abstract features; however, deep learning offers much better and more accurate results than traditional image processing methods. In this work, a methodology is proposed that supports the identification and recognition of emotions present in a person through facial images, which will be analyzed with an evolutionary convolutional neural network, in order to increase the reliability in the extraction of characteristics. Intuitively more complete and abstract images for their classification and identification of emotions. The proposed CNN to be modified performs feature extraction in three parallel stages. Each stage evolves with a genetic algorithm that in turn communicates with the other processes of the other stages based on the islands' communication paradigm.

References

[1] Avilés-Cruz, Carlos, Andrés Ferreyra-Ramírez, Arturo Zúñiga-López, and Juan Villegas-Cortéz. 2019. "Coarse-Fine Convolutional Deep-Learning Strategy for Human Activity Recognition." Sensors 19 (7): 1556. https://doi.org/10.3390/s19071556.

[2] Dang, Vu-Tuan, Hong-Quan Do, Viet-Vu Vu, and Byeongnam Yoon. 2021. "Facial Expression Recognition: A Survey and Its Applications." In 2021 23rd International Conference on Advanced Communication Technology (ICACT), 359–67. IEEE. https://doi.org/10.23919/ICACT51234.2021.9370369.

[3] Ko, Byoung. 2018. "A Brief Review of Facial Emotion Recognition Based on Visual Information." Sensors 18 (2): 401. https://doi.org/10.3390/s18020401.

[4] Villegas-Cortez, Juan, Carlos Avilés-Cruz, Arturo Zúñiga-López, Salomon Cordero-Sanchez, Francisco Fernández De Vega, and Francisco Chávez de la O. 2018. "Evolution of Statistical Descriptors for the Image Recognition of Natural Sceneries by Means of Genetic Programming for CBIR Improvement." In 2018 Seventeenth Mexican International Conference on Artificial Intelligence (MICAI), 45–50. https://doi.org/10.1109/MICAI46078.2018.00015.

Classification of falls events in the elderly using machine learning techniques

Arnoldo Díaz-Ramírez[a], Julia Diaz-Escobar[a], Verónica Quintero[a] and Rosendo Moncada[a]

[a]Department of Computer Systems
Tecnológico Nacional de México/IT de Mexicali
Av. Tecnológico s/n, Col. Elías Calles,21376, Mexicali, B.C, México
{adiaz, jdiaz, veronicaquintero}@itmexicali.edu.mx, rosendo.msa@gmail.com

Falls are of the main causes of injuries in older adults. In 2019, falls among adults ages 65 and up were the cause of over 34,000 deaths, being the leading cause of injury and death for the elderly [?]. If an older adult falls, the time elapsed between the fall event and the assistance to the patient is crucial since a rapid intervention has a great impact on the quality of life of the older adult. In contrast, if the assistance is not given timely, the patient may have serious and permanent injuries or even death.
In this paper, a model for the classification of falls in the elderly is introduced based on the Internet of Things paradigm and machine learning algorithms. The model uses a thermal sensor to monitor elderly activities. Three supervised machine learning models were selected for the fall and non-fall events classification, including a support vector machine, a random forest, and a convolutional neural network. For the training and testing stage, we utilized the publicly available falls dataset comprising 1253 thermal measure files divided into fall and non-fall classes. To achieve a high model performance, hyperparameter tuning was carried out. The performance results obtained by the models are presented employing repeated 3×10-fold cross-validation in terms of accuracy, balanced accuracy, and AUC-ROC. The obtained results show that the selected models successfully distinguish between fall and non-fall events. All three models obtained a performance above 95% for all metrics. Nevertheless, the random forest classifier achieved the best performance, up to 99% for all metrics.

Evaluation of Satisfaction, Personality, and Preferences Models applied in Project Portfolio Optimization

Jorge Castro-Rivera[a], María Lucila Morales-Rodríguez[b]
Nelson Rangel-Valdez[c], Claudia Guadalupe Gómez-Santillán[d]

[a,b,c,d]División de Estudios de Posgrado
TecNM/Instituto Tecnológico de Ciudad Madero
Juventino Rosas, Jesús Urueta esquina, Los Mangos, C.P. 89460, Cd. Madero, Tamps.
{D10070283, lucila.mr, nelson.rv, claudia.gs}@cdmadero.tecnm.mx

In decision-making, Satisfaction, Personality, and Preferences (SPP) play an important role in decision-makers (DM, Decision-Maker), especially for the development of cognitive models of virtual agents. For example, in a food context, preferences reflect ideal foods according to the DM's taste, satisfaction reflects acceptance concerning food, and personality influences food taste and acceptance. There are research works where multi-criteria approaches (outranking) are used to model preferences, managing to reflect the interests of the DM [1]. Similarly, there are research works that have managed to create computable models that integrate personality traits and satisfaction metrics to influence the DM's decision-making [2]. However, related works offer a limited characterization of the impact of SPP models on optimization problems. This is because only a single performance metric is analyzed using a small number of instances and model combinations. This makes it difficult to observe the relevance of the role played by SPP in the decision-making process. In this work, a methodology is proposed for the analysis of the impact on cognitive models of SPP in optimization and decision-making contexts. The proposed methodology consists of the following steps: (1) select case study, (2) generate random instances, (3) define reference set per instance, (4) select optimization algorithm, (5) implement cognitive model, (6) define performance metrics based on satisfaction and Euclidean distance, and (7) define the experimental design and implement it. This research considers a case study the Project Portfolio Problem (PPP) with a focus on the purchase of food. The cognitive models to be considered involve all possible configurations of the models considered. The main contributions are the methodology, the instance generator for the PPP, the performance metrics, and the characterization of advantages/disadvantages of SPP on decision making. As part of the results, it is expected to reduce the gap between the emulated decisions of an agent and the real decisions of the emulated DM.

References

[1] E. Fernández, E. López, F. López, and C. A. Coello-Coello, "Increasing Selective Pressure Toward the Best Compromise in Evolutionary Multiobjective Optimization: the Extended NOSGA Method," *Journal Information Sciences: an International Journal*, vol. 181, no. 1, pp. 44–56, 2011. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0020025510004354

[2] J. Castro-Rivera, M. L. Morales-Rodríguez, N. Rangel-Valdez, C. Gómez-Santillán, and L. Aguilera-Vázquez, "Modeling preferences through personality and satisfaction to guide the decision making of a virtual agent," *Axioms*, vol. 11, no. 5, 2022. [Online]. Available: https://www.mdpi.com/2075-1680/11/5/232

Multi-Objective Optimization of Microalgae Metabolism

Mónica Fabiola Briones Báez [a], Luciano Aguilera Vazquez [a]Nelson Rangel-Valdez [a]

[a]División de Estudios de Postgrado e Investigación
TECNM/Instituto Tecnológico de Ciudad Madero
Av. 1o. de Mayo esq. Sor Juana Inés de la Cruz s/n Col. Los Mangos, Cd. Madero, Tamaulipas, México
D00070469@cdmadero.tecnm.mx
luciano.av@cdmadero.tecnm.mx
nelson.rv@cdmadero.tecnm.mx

A metabolic network consists of a set of interconnected chemical reactions called metabolic pathways, this is represented mathematically as a stoichiometric matrix S of size *(m* x *n)*; an alternative representation can be achieved with the use of graphs, where each row represents a single metabolite existing in the metabolic system for a system with m components and each column represents one of the n reactions. The graphic scheme of a metabolism also reaction's forks into two or more pathways, which in a complex network might involve a large number of them. However, the yield of a particular product only depends on the regulation of the fluxes of a small subset of reactions that comprise the metabolic network usually distinguished as the main nodes. In some cases, the regulation and control of metabolic fluxes can be linked with the activation/deactivation of specific enzymes, in others it can also be linked to the use of specific growing conditions such as the type of nutrients where a particular pathway can be favoured.

So far, different computational techniques have been used for metabolic networks' analysis. For example, FBA (Flux Balance Analysis), identifies optimal sets of fluxes values in situations involving the study of a single objective. Alternatively, metaheuristics are used in scenarios involving the optimization of multiple bioproducts. However, those strategies still do not consider the regulation of fluxes in a metabolic network, a condition that naturally occurs in the study of metabolism, and its inclusion supports engineering in the search for better explanations of metabolic behavior.

For these reasons, the present research proposes a novel metabolic network representation that delves into the use of restrictions inside metabolic networks. The main purpose of the research is to reflect the control and regulation of metabolism so that it is possible to adapt it to realistic environmental conditions. The model is solved using NSGAII with the expectation of reaching a controlled way of adding conditions in the behavior of the fluxes to guide the production of metabolites of interest.

References

[1] George Stephanopoulos, et al., Metabolic engineering: principles and methodologies, 1998

[2] Zuñiga, C., et al., Genome-scale metabolic model for the green alga Chlorella vulgaris UTEX 395 accurately predicts phenotypes under autotrophic, heterotrophic, and mixotrophic growth conditions, Plant physiology, 172(1), 589-602.

Towards chemical forecasts guided by negotiating software agents

José Ángel Villarreal-Hernández[a], María Lucila Morales-Rodríguez[b], Nelson Rangel-Valdez[c]

[a,b,c]Departmento de Estudios de Posgrado e Investigación
TecNM/Instituto Tecnológico de Ciudad Madero.
Ciudad Madero, Tamaulipas, México.
{[a]D10071371, [b]lucila.mr, [c] nelson.rv}@cdmadero.tecnm.mx

Chemical prediction is a problem that arises from the need for chemistry professionals to identify the working configurations of interest before performing experiments in the laboratory, helping to reduce time and material costs. However, chemistry professionals are faced with a scenario in which there is a large number of prediction tools to resort to and these require a configuration of parameters to operate. In general, chemical professionals do not have the expertise to make these decisions. The automatic configuration of these parameters can be a contribution to them. Intelligent software agents have been involved in supporting similar tasks, emulating the behavior of the human decision-maker. Additionally, intelligent agents can test different combinations in a simulated way before making a decision. The agent paradigm can be a means to address the algorithm configuration problem. This would be beneficial to professionals in chemistry among others.

Data mining and machine learning techniques have been used in chemical prediction. According to Hao Li et al. [3], data mining techniques successfully predict the trend of data and have been used to solve chemical process, material process, and engineering process problems. Like other types of algorithms and computational techniques, data mining and machine learning applications require the user to define specific settings for running the algorithm to produce the expected results. For example, in the KNN algorithm, it is necessary to indicate the K value of close neighbors to consider classifying an element. It is known that the performance of the algorithms is related to their parametric configuration, and in the literature, there are several techniques to treat its adjustment [2]. The parameters guide the behavior of the algorithms, and it is usually the researcher who estimates which parameter settings have the best chance of producing the desired effects. However, their estimates may not be optimal or have an unclear degree of reliability.

When a computer expert is choosing how to deal with a prediction problem, they make compromises as 1) what algorithm is convenient to use, 2) what parameter configuration to use in the algorithm and 3) what modifications are necessary for the data body for a correct operation. Consider that these compromises points can affect each other. For example, deciding on a specific algorithm can lead to applying changes to the data body, just as characteristics of the data body could lead us to choose one algorithm over another. Thus, to effectively locate the best option among the compromises, they must be balanced. With the agent paradigm, we can propose a way to automate this balance, through the interactive process of negotiation. Negotiation is a process that tries to reach an agreement. This agreement can be the solution to situations in which it is necessary to reach commercial deals, resolve conflicts, or form alliances [1]. In this context, negotiators are deliberative software agents that represent each compromise point. During the negotiation process, each party would propose to the others a set of options that maximize their own performance. But for a proposal to be accepted by the other parties, it must be good enough for its objectives and that is where the balance is produced.

The paradigm of agents is considered to offer beneficial characteristics for problem-solving: agents can work independently applying different computational techniques and, endowed with attributes such as exploration and experience, they can evolve their way of proceeding to optimize the results provided. Given the treated elements, we can visualize a multi-agent system that decides which algorithm to use, with which parametric configuration, and how to modify the data body, all this for a specified chemical prediction scenario.

References

[1] Tim Baarslag. *What to bid and when to stop.* Technische Universiteit Delft, Dutch research school for Information and Knowledge Systems, 2014. OCLC: 905871136.

[2] Mauro Birattari. *The problem of tuning metaheuristics as seen from a machine learning perspective*. PhD thesis, Universite Libre de Bruxelles, August 2005.

[3] Hao Li, Zhien Zhang, and Zhe-Ze Zhao. Data-Mining for Processes in Chemistry, Materials, and Engineering. *Processes*, 7(3):151, March 2019.

Knowledge discovery through an archimedean compensatory fuzzy logic neural network.

Carlos Eric Llorente-Peralta[1*], Jose Fernando Padron-Tristan[1], Laura Cruz-Reyes[2], Rafael Alejandro Espin-Andrade[3],

[1] Division of Postgraduate Studies and Research, National Institute of Mexico/Technological Institute of Tijuana, Unidad Tomás Aquino, Calzada Tecnológico S/N Esq. Av. Castillo de Chapultepec y calle Cuauhtemotzin, Fracc. Tomás Aquino C.P. 22414. [2] Division of Postgraduate Studies and Research, National Institute of Mexico/Technological Institute Ciudad Madero, Av. 1o. de Mayo esq. Sor Juana Inés de la Cruz s/n Col. Los Mangos, Cd. Madero, Tamaulipas, México.
[3]Research and Postgraduate Directorate, Autonomous University of Coahuila
* E-mail: ceric.lop@gmail.com

Key words: Neural networks, Archimedean compensatory fuzzy logic, predicate logic, knowledge discovery.

The information generated in the world is growing at a fast pace, so it is estimated that the amount of existing data is doubled approximately every 20 months. However, having this amount of data does not only provide any relevant information. That is why, within the area of computational sciences, knowledge discovery (KD) arises, which can be defined as "the process of non-trivial extraction of implicit, previously unknown and potentially useful information from data." [1, 2]. To this end, various models have been created whose purpose is to discover existing patterns, rules, and relationships in data warehouses. Within these models, the use of artificial neural networks (ANN) emerges as a promising alternative in these processes; due to the way it operates, an ANN can be trained to learn or discover existing knowledge in the data. Where its main strength is the ability to adapt and self-organize [3]. In the following document, a new ANN model with compensatory Archimedean fuzzy logic (ACFL) is proposed, which has been defined as a compensatory Archimedean fuzzy logic hybrid neural network (NN-ACFL). neuronal of the feed-forward type, however, the concepts that allow training is replaced by concepts of the ACFL, thus allowing the generation of a network that avoids the use of extra logical resources in the learning process. Among the advantages of a transdisciplinary logical theory such as the ACFL is the use of a generalized continuous linguistic variable, through which it is possible to generate a family of membership functions, which adapt to the information, allowing the generation of the function that best adapts to the discovery process, on the other hand, the ACFL, allows interpreting the results and presenting them in a language close to nature through which it is possible to justify the result of the information [4] [5, 6]. In addition, the proposed network generates a predicate in a disjunctive normal form (DNF) which is the data search and optimization model; the advantage of the use of logical predicates is that these structures allow dealing with various data mining problems, as well as it is possible represent them through different structures such as graphs, trees, among others , Thus, constructing these structures also allows their interpretation in natural language in a relatively simple way.. For the optimization process, the Backpropagation algorithm is used, which allows for optimizing the concepts of the ACFL, improving the results in each epoch of the NN-ACFLAmong the results obtained, we find that, as expected, the ACFL allows the network to work in a similar way to how it is done classically. However, due to its characteristics, each concept within the network has a different interpretation. Linguistica, which not only allows finding a result through the training processes of an NN, but also expressing the results in a language close to natural. Thus, it is also observed that the model offers competitive results concerning the discovery of knowledge and allows the construction of classification and support models in decision-making processes.

Bibliografía

[1] W. Frawley, G. Piatetsky-Shapiro and C. Matheus, "Knowledge Discovery in Databases: An Overview.," *AI Magazine,* p. 57, 1992.

[2] M. Pazzani, "Knowledge discovery from data," *IEEE Intelligent Systems,* vol. 15, no. 2, p. 2, 2000.

[3] F. LiMin , «Knowledge discovery based on neural networks,» *Communications of the ACM,* vol. 42, nº 11, p. 3, 1999.

[4] R. A. Espin-Andrade, E. Gonzalez, W. Pedrycz y E. Fernandez, «An Interpretable Logical Theory: The case of Compensatory Fuzzy Logic,» *International Journal of Computational Intelligence Systems,* vol. 9, nº 4, p. 14, 2016.

[5] R. A. Espin-Andrade, . E. Gonzalez, . W. Pedrycz y E. Fernandez-Gonzalez, «Archimedean-Compensatory Fuzzy Logic Systems,» *International Journal of Computational Intelligence Systems,* vol. 8, nº 2, p. 8, 2015.

Consistent Conjectural Variations Equilibrium for a Financial Model

José Guadalupe Flores-Muñiz[a], Nancy Solis-García[a], Nataliya Kalashnykova[a], and Viacheslav Kalashnikov[b]

[a]Department of Physics and Mathematics
Universidad Autónoma de Nuevo León
San Nicolás de los Garza, Nuevo León, Mexico

[b]Department of Systems and Industrial Engineering
Instituto Tecnológico y de Estudios Superiores de Monterrey
Monterrey, Nuevo León, Mexico

We consider a general model of financial flows and prices with multiple sectors and instruments. Each sector optimizes the composition of assets and liabilities in its portfolio, whose utility is given by a quadratic function constrained to satisfy the accounting identity that appears in flow-of-funds accounts and the equilibrium conditions that guarantee market clearance. To define the financial equilibrium, we make use of the concept of conjectural variations, in which each sector conjectures the possible dependence of the instruments' prices upon its portfolio structure. The problem is modeled as a continuous two-stage game. In the first stage, the set of strategies for each sector consists of its possible conjectures about its influence on the prices. The second stage is the financial model's equilibrium problem where, according to the conjectures selected in the first stage, each sector decides its portfolio composition. Each sector aims to minimize the risks, while at the same time maximizing the value of its assets and minimizing the value of its liabilities.

The importance of WRF parameterization for wind ramp prediction

García-Rodríguez I.[a], Cadenas-Calderón E.[a], Pacheco-Ibarra J.[a]
Campos-Amezcua R.[b], Bernabé-Morales M.[a]

[a]Faculty of Mechanical Engineering
Universidad Michoacana de San Nicolás de Hidalgo
Santiago Tapia 403, Centro, 58000 Morelia, Mich.
9602418a@umich.mx, ecadenas@umich.mx, jpacheco@umich.mx, 091558d@umich.mx

[b]National Center for Technological Research and Development
Tecnológico Nacional de México
Interior Internado Palmira S/N, Col. Palmira, Cuernavaca 62490, México
rafael.ca@cenidet.tecnm.mx

The following numerical model is presented using the WRF tool for a 24-hour horizon. This work aims to demonstrate the convenience of identifying the best parametric configuration of the model by comparing it with the configuration that the software proposes by default. The parameterizations incorporate into the model different physical phenomena approximations that take place on a smaller scale than the model's time step integration, such as microphysics, surface layer, soil surface, etc. [1]. The above was used to identify qualitatively and quantitatively phenomena known as wind ramps, which must be predicted in advance so that the wind energy system can handle the abrupt fluctuation in energy [2]. To verify the results, we can cite historical measurements of the place, La Mata, Oaxaca. The results show that the proper management of the WRF parameterizations makes the wind ramps could be identified with greater precision.

References

[1] M. Castro, R. González y A. Portela, Parametrización de la capa límite atmosférica en los modelos numéricos de pronóstico meteorológico, Física de la tierra, num 3,1991, Ed. Universidad Complutense Madrid, pp. 103-138.

[2] Y. Zhang, S. Gao, J. Han, and M. Ban, "Wind speed prediction research considering wind speed ramp and residual distribution,", 2019 IEEE Acces, vol. 7, pp. 131873-131887, doi: 10.1109/ACCESS.2019.2940897

Intelligent Forecasting Methods for COVID-19, MLP-SVRES

Gerardo de Jesús Martínez Neri[a]**,Juan Frausto-Solis**[a*]**, Juan J. González-Barbosa**[a], **Lucía J. Hernández-González**[a]
[a]División de Estudios de Postgrado e Investigación
TECNM/INSTITUTO TECNOLÓGICO DE CIUDAD MADERO
Av. 1o de Mayo esq. Sor Juana Ines de la Cruz s/n Col. Los Mangos C.P.89440
* juan.frausto@gmail.com

Abstract

Since 2019, many people in Mexico have become ill with coronavirus disease (COVID -19), which is caused by the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-Cov2) virus. For this reason, the health protocols established throughout the country to combat the spread of the virus must be adapted according to the constant changes in contagion. The fundamental objective of a forecast is to reduce the range of uncertainty within which decisions are made that affect the future and, therefore,with it all parties involved. The help of such forecasts in determining new measures in the fight against COVID is expressed. In this work, we present a model for a 10-day forecast of the infection cases of COVID -19 in Mexico called Multi-layer Perceptron - Support Vector Regression with Exponential Smoothing (MLP-SVRES). Support Vector Regression and Multi-layer Perceptron are used as forecast methods. Exponential Smoothing is used as an enhancement method. The training period is defined from August 1 to October 31, 2021. The results show that applying the exponential smoothing method drastically reduces the error which we measure with the MAPE metric.

Keywords: Support Vector Regression, Multi-Layer Perceptron, Exponential Smoothing, SARS-CoV-2, COVID-19, forecasting.

1 Introduction

The COVID-19 pandemic has affected the lives of millions of people around the globe. In addition to the serious health consequences, it has devastated our goals, family dynamics, work roles, and economic stability. In this sense, it constitutes an unprecedented global crisis. To forecast the number of confirmed cases, deaths, and fatalities, different models are built and optimized, which are adjusted with the data published by the Ministry of Health in Mexico, and the data for the U.S.
Once the models are built, those that provide the lowest error of fit with respect to the historical data are selected [1]. Subsequently, the selected models are estimated with historical data. Finally, to forecast the number of confirmed cases, deaths, and hospitalizations for the following days, the average value of the prediction for each day is calculated, as well as its estimation error with different confidence levels.

2 Proposed methodology

In this proposed method we entered the official data obtained from the Open Data of the General Directorate of Epidemiology of Mexico [2]. Figure 1 shows the proposed methodology called MLP-SVRES. We work with a time series of confirmed cases of COVID-19 from August 1 to October 31 2021 in Mexico. The data were divided into training, validation, and test with the following percentages 77%, 11%, and 11%.
The methods used in this work are a multi-layer perceptron (MLP) and a Support Vector Regression (SVR). SVR is used as a method for forecasting daily cases of COVID-19. Likewise, MLP is used for this purpose. Both methods forecast 10 days for confirmed cases of COVID-19. SVR and MLP forecasts are subjected to an Exponential Smoothing (ES) enhancement process. Enhancement ES forecasts are evaluated with the Mean Absolute Percentage Error (MAPE) metric.

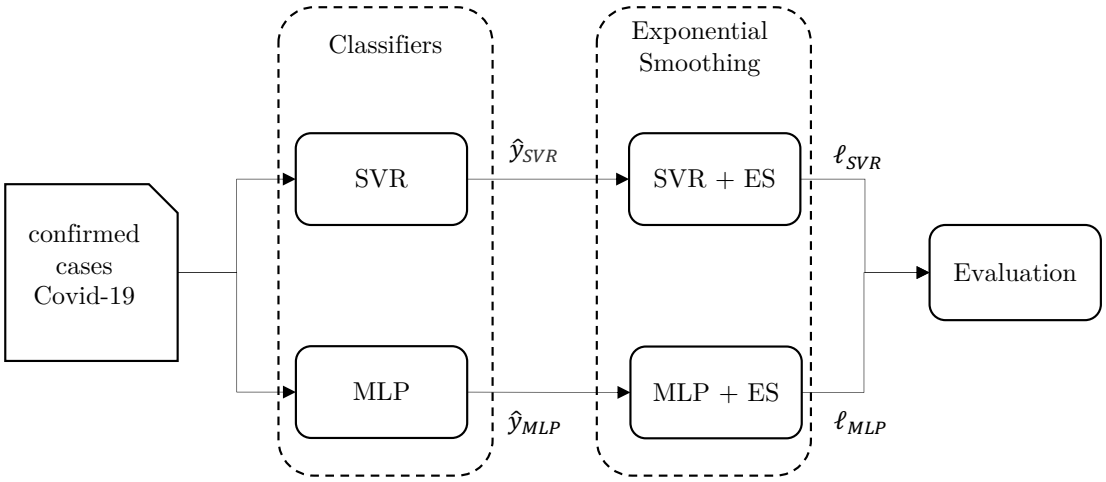

Figure 1: MLP-SVRES methodology to forecasting daily cases of COVID-19 in Mexico using SVM and MLP as forecast methods and exponential smoothing (ES) as enhancement method.

3 Results

The results obtained are demonstrating the importance of using intelligent methods for forecasting COVID-19 in Mexico. Table 1 shows that the MAPE of SVR with ES is approximately 20% better than that obtained by MLP with ES.

Table 1. Performance comparison of classifiers using the MAPE metric

Classifiers	MAPE
SVR + ES	**18.34**
MLP + ES	38.91

4 Conclusion

In this work, we applied two methods SVR and MLP to forecasting daily confirmed cases of COVID-19 for Mexico. We observe that the error from enhanced forests in MLP is more pronounced in comparison to SVR. The results show that SVR is 20% better than MLP.

References

[1] Bianco, M. J., Cruz, P. D., Fraquelli, A. D., Gache, A. L. (2020). Modelo epidemiológico SIR: Una aplicación de las ecuaciones diferenciales al SARS-CoV-2 (COVID-19). Revista de Investigación en Modelos Matemáticos aplicados a la Gestión y la Economía, 7(1), 16-38.

[2] Secretaría de Salud en Mexico. (2015, 20 febrero). Datos Abiertos Dirección General de Epidemiología. Datos Abiertos Dirección General de Epidemiología. Recuperado 1 de noviembre de 2021, de https://www.gob.mx/salud/documentos/datos-abiertos-152127

Identification of the spread of COVID-19 in Mexico, using temporal complex networks and optimization approaches

Edwin Montes-Orozco[1], Roman Anselmo Mora-Gutiérrez[2], Christian Lizbeth Noguez-Moreno[3]*, Ángel David Téllez-Macías[3], Gilberto Sinuhé Torres-Cockrell[4] and Miguel Ángel Fernández-Romero[5]
[1]Departamento de Matemáticas Aplicadas y Sistemas. Universidad Autónoma Metropolitana Cuajimalpa.
[1]Departamento de Sistemas. Universidad Autónoma Metropolitana Azcapotzalco.
[3]Departamento de Tecnologías de la Información. Universidad Autónoma Metropolitana Cuajimalpa.
[4]Posgrado en Ciencias y Tecnologías de la Información. Universidad Autónoma Metropolitana Iztapalapa.
[5]Independent Consultor

Abstract

In this work, we present the characterization of the spread of the disease known as COVID-19 in Mexico. This characterization is developed on a temporary complex network considering periods of 3 months from the first quarter of 2020 to the second quarter of 2022; therefore, the analysis is carried out in 10 different time periods, where each node represents a municipality in Mexico and each link takes weight based on: number of confirmed cases, number of possible cases and number of deaths, related to COVID-19 as well as adjacencies municipalities, number of inhabitants in hospital infrastructure. On the other hand, to verify which municipalities have been the most spreaders of COVID-19, we use the approach of inverse percolation in complex networks by resolution of the vertex separator problem through a genetic algorithm.

*Corresponding author.

Benchmarking a CUDA-based implementation of genetic programming using SRBench

Luis A. Cardenas Florido, Leonardo Trujillo, Daniel E. Hernández

Posgrado en Ciencias de la Ingeniería
Tecnológico Nacional de México/IT de Tijuana
Tijuana, BC, Mexico
armando.cardenas@gmail.com, leonardo.trujillo@tectijuana.edu.mx, daniel.hernandezm@tectijuana.edu.mx

Genetic programming (GP) is one of the leading methods in the problem domain of symbolic regression, a supervised learning problem where the goal is to find a symbolic model that approximates the target variable in the training data. Unlike black-box models like neural networks or ensembles of regression trees, symbolic regression models are amenable to human interpretation and understanding. Recently, a comprehensive benchmark suite for symbolic regression has been made public, called SRBench [1]. The suite allows researchers to rigorously evaluate and compare their symbolic regression algorithms on a wide variety of synthetic and real-world problems, including several state-of-the-art techniques of both GP and non-GP methods. However, among the considered methods none of the GP-based approaches exploit GPU-based computation, a cornerstone of modern machine learning implementations. This work presents the first evaluation through SR-Bench of a CUDA-based GP algorithm, based on our previous CUDA implementation of geometric semantic GP [2]. Several variants of our algorithm are evaluated, and results show that the method is competitive with other basic GP systems in terms of quality, but outperforms them based on efficiency and model size. The former result was expected given the use of CUDA and GPUs in our implementation, but the latter was unexpected and is probably due to the stack-based representation used for the evolved individuals [2].

References

[1] W. La Cava, et al., Contemporary Symbolic Regression Methods and their Relative Performance, Neurips Track on Datasets and Benchmarks, 2021.

[2] L. trujillo, et al., GSGP-CUDA — A CUDA framework for Geometric Semantic Genetic Programming, SoftwareX, Volume 18, June 2022, 101085.

Selection methods applied to FCTA* forecasting ensemble methodology

Lemuel Rodriguez-Moya[1], Juan Frausto-Solís[1], Juan J. González-Barbosa[1], Guadalupe. Castilla-Valdez[1]

[1] Tecnológico Nacional de México/IT Ciudad Madero, México

Abstract
The combination of forecasting methods is widely applied technique in recent years. FCTA is a new ensemble method which is too robust to combine classical methods, as ARIMA and Simple Exponential Smoothing, and intelligent methods as the classical Neural Networks, Support Vector Regression, deep learning, and even other hybrid methods. FCTA always obtains a best forecasting that the individual methods in the ensemble. Nevertheless, if the selected methods used in FCTA are not the most precise in the area, the prediction obtained would not be as good as needed. In this paper we present a new ensemble methodology named FCTA* which enhance other ensemble forecasting methods. This new methodology used the classical B&G (Bates and Granger) to refine the weights obtained with FCTA. According to the experimentation with a dataset presented in the paper, the new methodology or FCTA* is in general better than FCTA and other ensemble methods.
Keywords: Combining Forecasting Methods, Time Series, Forecasting BIC, Threshold Accepting.

1. Introduction

Methods combined with FCTA methodology [1] always produce a better forecast than the best individual method. This methodology is applied to any area, and basically consists of the next steps: A) firstly, it selects the best forecast methods in the field; B) they test the selected methods in a first validation stage using metric forecasting errors; in this stage they estimate a first prediction using a set of weights of each individual forecasting error; C) the last estimation is refined by a especial threshold heuristic method [2] which minimizes the forecasting error of the last stage. In this stage the best weights of the individual methods are determined. However, the used methods in FCTA could not be the best ones. B&G methodology [3] established that the best combination of individual methods is obtained using an optimization model and a heuristic based on the total variance of the entire set of individual method. In the new methodology FCTA*, the selection of individual methods is enhanced by a heuristic based on the BIC metric [4] and is refined with a B&G method. We also used a Threshold accepting algorithm to optimize the weights used for each individual forecast. In the final paper we compare the new ensemble FCTA* methodology with other ensemble methodologies of the literature.

2. Method

FCTA* methodology [1] is performed using selection strategies based on the classical BIC metrics [4]; they are used in a new function which determines when an individual forecasting method may be included in the ensemble. Then a threshold algorithm [2] is used to determine the best weights for each individual method. Besides, a B&G method [4] is applied to check to refine the set of methods used by FCTA. There are successful strategies such as Bates and

Granger (B&G). This as well as FCTA set weights for all individual methods. B&G determines weights using the correlation between the variance of the forecasting methods in the combination.

3. Results

The final paper shows some results comparing FCTA* with FCTA and other ensemble methods. We will show than the FCTA* obtains best results in most of the cases.

4. Conclusions

This paper presents a new FCTA* ensemble forecasting methodology. It includes a new selection of individual forecasting methods. As is shown in the paper, the forecasting results obtained with FCTA* are better than other ensemble methodologies.

4. References

[1] J. Frausto-Solis, L. Rodriguez-Moya, J. González-Barbosa, G. Castilla-Valdez, and M. Ponce-Flores, "FCTA: A Forecasting Combined Methodology with a Threshold Accepting Approach," *Math. Probl. Eng.*, vol. 2022, 2022, doi: 10.1155/2022/6206037.

[2] G. Dueck and T. Scheuer, "Threshold accepting: A general purpose optimization algorithm appearing superior to simulated annealing," *J. Comput. Phys.*, vol. 90, no. 1, pp. 161–175, 1990, doi: 10.1016/0021-9991(90)90201-B.

[3] J. M. Bates ; C. W. Granger, "Combination of Forecast," *Oper. Res. Q.*, vol. 20, no. 4, pp. 451–468, 1969, doi: 10.2307/3008764.

[4] S. Tofighy, A. A. Rahmanian, and M. Ghobaei-Arani, "An ensemble CPU load prediction algorithm using a Bayesian information criterion and smooth filters in a cloud computing environment," *Softw. - Pract. Exp.*, vol. 48, no. 12, pp. 2257–2277, 2018, doi: 10.1002/spe.2641.

Improving genetic programming based approximations of the hypervolume indicator

Cristian Sandoval[a], Leonardo Trujillo[a], Luis C. González[b]

[a]Posgrado en Ciencias de la Ingeniería
Tecnológico Nacional de México/IT de Tijuana
Tijuana, BC, Mexico
sr_cristian@outlook.es, leonardo.trujillo@tectijuana.edu.mx

[b]School of Engineering
Autonomous University of Chihuahua
Chihuahua, Chihuahua, Mexico
lcgonzalez@uach.mx

In multi-objective optimization, particularly when solved by evolutionary multi-objective algorithms (EMOAs), the hypervolume (HV) is the most widely used performance indicator, used to evaluate the quality of Pareto front approximations. Moreover, the HV can be used to turn a multi-objective problem into a single objective problem, by posing search strategies that directly optimize the HV values, such as indicator based MOEAs like the popular SMS-EMOA. An important drawback, however, when using the HV indicator is it's computational cost, which makes it difficult to use with four or five objectives, and practically impossible for more. For this reason many authors have developed approximation strategies, such as using Monte Carlo strategies or the achievement scalarizing function method. In our previous work, we presented approximating functions that were generated automatically, by posing a supervised learning problem and solving it using Genetic Programming (GP) [1]. The evolved models were highly accurate and much more efficient than any previous approximation methods, and were shown to be capable of guiding an indicator-based EMOA to competitive Pareto front approximations on standard benchmarks with 3, 4 and 5 objectives. The results presented in [1], however, do have some drawbacks. First, the evolved models were only applicable to a certain number of objectives; i.e., a model evolved for problems with n objectives could not be used on a problem with m number of objectives. Second, when used to guide SMS-EMOA, the models had difficulty once the population was close to converging on some problems; i.e., the model had trouble estimating the HV contributions of solutions that were part of a large set of non-dominated solutions that were close to the true Pareto set. Finally, the models where not tested on many-objective problems, with more than five objectives. This work will discuss the first steps towards addressing these issues. In order, we present a first an initial attempt to generate models that can generate to different numbers of objectives, heuristics to deal with the convergence problem, and initial results on problems with as many as 10 objectives.

References

[1] Cristian Sandoval, et al., Towards fast approximations for the hypervolume indicator for multi-objective optimization problems by Genetic Programming, Applied Soft Computing, Volume 125, 2022, 109103.

Xalapa

NEO 2022 Organizing Committee

- **General Chairs**
  Marcela Quiroz, Universidad Veracruzana, México

- **Publicity Chair**
  Nelson Rangel Valdez, Instituto Tecnológico de Cd. Madero, México

- **Students Chair**
  Daniel Hernández, Instituto Tecnológico de Tijuana, México

- **Series Chairs**
  Oliver Schuetze, Cinvestav-IPN, Mexico
  Leonardo Trujillo, Instituto Tecnológico de Tijuana, México

Index

J

K

L

M

N

O

P

Q

R

S

T

U

V

Y

Z