MODERN SOFT COMPUTING TECHNIQUES AND THEIR APPLICATIONS IN CIVIL ENGINEERING

Ali Danandeh Mehr

[1]Department of Civil Engineering, Antalya Bilim University, Antalya, Turkey,
ali.danandeh@antalya.edu.tr

Keywords: Soft Computing, Genetic Programming, Civil Engineering,

ABSTRACT:

Soft computing (SC) is a group of techniques and methodologies applied to solve a wide range of problems spread in several areas of science. The aim is to exploit tolerance for imprecision, uncertainty and partial truth to achieve tractability, robustness and low solution cost. Some of the well-stablished methodologies include Neural Networks, Evolutionary Computing, Fuzzy Logic, Rough Sets, Decision Trees, etc. The focus of this paper is to briefly review the fundamentals of modern genetic programming (GP) techniques and their recent applications in various disciplines of civil engineering, particularly in hydro-environmental studies.

1. INTRODUCTION

Over the past few decades, there have been considerable research on soft computing (SC) that has increased the human awareness in complicated issues and engineering problems. Generally, it defined as the use of inexact solutions to computationally hard tasks. Fuzzy Logic, Artificial Neural Networks, Evolutionary Computing, and decision trees are of well-known SC techniques that have been used extensively to solve a wide range of engineering problems. From the type of problem point of view, SC techniques have been used mainly to discover mathematical relationships between the empirically observed variables, which is called symbolic system identification or black box modeling. Once a model is discovered and verified, it can be used to predict future values of the state variables of the system.

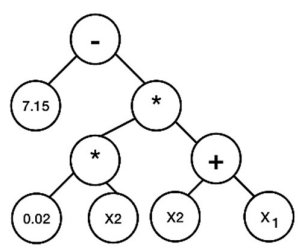
Genetic programming (GP) is a relatively new SC technique from evolutionary computing family. It is an extension of the well-stablished genetic algorithm in which the genetic population consists of computer programs. Each program is compositions of primitive functions and terminals. It was applied to different tasks in Civil Engineering such as structural optimization, design problems, classification, regression problems, and time series forecasting/modeling. This paper provides a general background to the GP and its advancements that is expanded in subsequent sections to demonstrate how its variants were effectively utilized in the recent (2017-2019) studies of different civil engineering disciplines.

2. FUNDAMENTALS OF GENETIC PROGRAMMING (GP)

GP is a domain-independent, problem-solving approach in which computer programs are evolved to find solutions to problems. The solution technique is based on the Darwinian principle of 'survival of the fittest' [1]. The programs are typically characterized by a tree structure known as genome. Evolutionary operators act on the genomes. This is the original structure of GP; however, there are other GP structures where computer programs are not evolved necessarily based on tree-shaped genomes. The new structures have been mainly suggested to speed up the run time and achieve more proper solutions. In this section, fundamentals of original GP (hereafter classic GP) and its advancements including multigene genetic programing (MGGP), Linear GP (LGP), and Gene Expression Programming (GEP are briefly described. For details about each approach the interested reader is referred to [2].

2.1. Classic GP

As previously mentioned, GP is an evolutionary algorithm that generates and optimizes computer programs (called individuals) to solve a problem. It has been widely used to solve symbolic regression problems. However, it is also able to solve binary or multiclass classification problems [3, 4]. The main reason for its great success over time is it's explicitly and ability to evolve mathematical expressions that can be easily analyzed, validated and applied in practice. In GP, the main operators that are used to convert the initial population of individuals (i.e., random programs) to a set of preferred solutions include crossover, mutation, and reproduction. These operators act on a single or a pair of tree shape individuals called the GP tree. GP trees can be commonly expressed by mathematical equations. For example, mathematical representations of a GP tree was illustrated in Figure 1.



$$y = 7.15 - 0.02\ x_2(x_2 + x_1)$$

Figure 1. Tree and mathematical representations of a GP program.

Figure 2 shows a standard flowchart of the main steps of GP. By crossover, some nodes/branches of a pair of GP trees are exchanged. It cause the algorithm to produce offspring closer to the targets. Reproduction is the direct transfer of the best tree to a new population of programs, and the mutation is the alteration of a node or branch in a single GP tree. Each node in a tree can adopt a function or terminal variables such as x_1 and x_2 in Figure 1. Some of the main issues in the GP-based modeling is the selection of a set of appropriate functions, input variables, and maximum depth (also referred to as height) of GP trees. Wisdom decision for these issues not only supports the algorithm to reach more accurate solutions but also helps it to avoid overfitting and complicated solutions.

2.2. MGGP

MGGP is a robust variant of GP that linearly integrates low-depth GP trees to increase the fitness of GP solutions. In MGGP, the target variable is calculated by the weighted sum of each gene in the multigene chromosome with a constant value called noise. Thus, an MGGP simplified model can be mathematically expressed as

$$\hat{y} = c_0 + c_1 \times \text{Gene } 1 + c_2 \times \text{Gene } 2 + \ldots + c_i \times \text{Gene } i \quad (1)$$

where the gene i is mathematical expression of standard GP for the ith gene evolved for the input parameters, c_0 is irregular term (noise) and $c_1, c_2, \ldots, c_i$ are the gene weights (i.e. regression coefficients) that are generally determined by the ordinary least squares method. For example, the pseudo-linear MGGP model shown in Figure 3 represents the output variable $\hat{y}$ as a combination of three standard GP trees comprising two input variables x_1 and x_2.

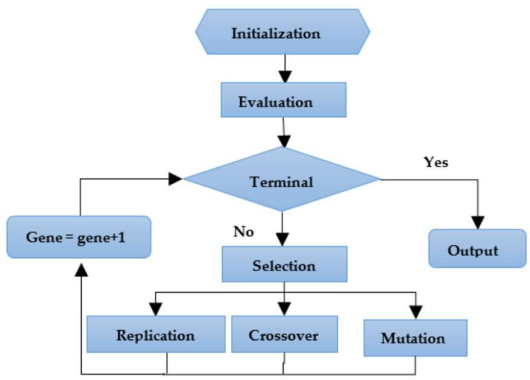


Figure 2. Flowchart of classic GP [5].

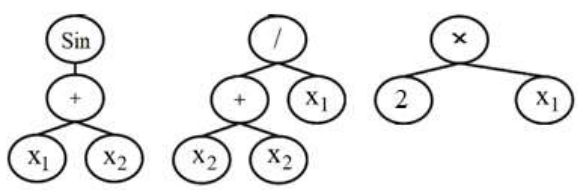


Figure 3. An example of a multigene individual involving three genes

The model is expressed in Equation 2.

$$\hat{y} = c_0 + c_1(\sin(x_1 + x_2) + c_2((x_2 + x_2)/x_1) + c_3(2 \times x_1) \quad (2)$$

2.3. Linear GP

LGP is another form of GP in which genetic operators act on a linear— not tree-based — GP (Figure 3). In other words, the program is expressed in a one-dimensional row of functions, terminals and constants. In Figure 4, each of instructions include a function that accept a minimum number of constants or memory variables with initial value of zero, called register (r). The r value consecutively changes at each instruction and the result is equal to the r value in the last instruction.

Instruction 1	Instruction 2	...	Instruction N

Figure 4 LGP representation with N instructions.

2.4. Gene Expression Programming

GEP (Ferreira 2002) is a type of GP in which the output variable is computed by linking some sub-expression trees using algebraic or Boolean functions. Indeed, GEP uses the power of small GP trees to capture nonlinear behavior of a complex system via a mixing procedure. Despite using multi genes (i.e., sub expressions), GEP does not provide a multigene solution and the final solution is a single GP tree. For example, a GEP model shown in Figure 5 denotes the output variable $\hat{y}$ as a production of three sub-expression linked together with multiplication function.

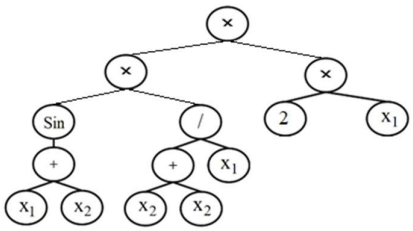


Figure 5 Example of the GEP individual involving three Sub-ETs linked by multiplication function

3. APPLICATIONS IN CIVIL ENGINEERING

This section summaries some of the recent applications of GP variants in Civil Engineering. The representative papers having wide range of applications are clustered in four domains of Water Resources Engineering (WRE), Geotechnical Engineering (GE), Structural and Earthquake Engineering (SEE), and Transportation and Traffic Engineering (TTE). A brief description of the reviewed papers together with their authors, publication year, and implemented GP variant are presented in Table 1.

4. CONCLUSION

GP is a powerful and flexible modeling technique that could be employed in various ways to solve complex problems in Civil Engineering. One of the most important points must be taken in to account during the application of GP to a specific problem is the dimensional characteristics of the inputs and target variables. For example, if a researcher aim to generate a formula for a process, dimensionally aware GP is preferable. Overfitting and complexity of solutions are of the main issues that must be considered during the modeling. Separation of the historical inputs into the training and unseen validation subsets is the widely used method to avoid overfitting. Use of low-depth GP trees, application of simple arithmetic functions, and generation under parsimony pressure are of the common ways that modelers utilized to control the complexity of solutions. Undoubtedly, GP is still a growing field of investigation, whose experts are still exploring its capabilities and limits.

Table 1: Summary of the surveyed papers that applied GP variants in Civil Engineering

Task	Authors	method	Description
WRE	Danandeh Mehr and Kahya 2017, [6]	MGGP	Streamflow forecasting using daily observed discharge
	Ravansalar et al. 2017, [7]	LGP	Time series forecasting using monthly data
	Heřmanovský et al. 2017, [8]	GP	Symbolic regression between rainfall, basin characteristics and runoff
	Chadalawada et al. 2017, [9]	GP	Evolution of optimal Tank model to represent Rainfall-Runoff processes
	Ghorbani et al. 2018, [10]	MGGP	Time series forecasting using daily data
	Mehdizadeh and Sales 2018, [11]	GEP	Time series forecasting using monthly data
	Danandeh Mehr 2018, [12]	GEP	month ahead streamflow forecasting in an intermittent stream
	Hadi and Tombul 2018, [13]	MGGP	Time series forecasting using monthly data
	Jamei and Ahmadianfar 2019, [14]	LGP	Prediction of scour depth at piers with debris accumulation
SEE	Abdollahzadeh et al. 2017, [15]	GEP	Prediction of compressive strength of high strength concrete
	Velay-Lizancos et al. 2017, [16]	OGP *	Estimating the compressive strength of eco concretes by combining non-destructive testing
	Vanneschi et al. 2018, [17]	NAGP **	High-performance concrete strength prediction
	Hodhod et al. 2018, [18]	MGGP	Prediction of creep compliance in concrete
	Asim et al. 2018, [19]	GP-Boost	Earthquake prediction using seismic indicators
	Qu et al. 2018, [20]	GP	Crack detection in concrete surface
GE	Goharzay et al. 2017, [21]	GEP	developing different models to evaluate the occurrence of soil liquefaction
	Soleimani et al. 2018, [22]	MGGP	Prediction of unconfined compressive strength of geo-polymer stabilized clayey soils
	Kurugodu et al. 2018, [23]	GP	Regression between strength improvement factor and soil moisture, density, and fiber content
	Puente et al. 2019, [24]	GP	Soil Erosion Estimation
	Ranasinghe et al. 2019, [25]	LGP	Predictions of effectiveness of rolling compaction with dynamic cone penetrometer test
TTE	Sahani et al. 2017, [26]	GP	classification of service measure at the roadside walking environment
	Chopra et al. 2018, [27]	GP	Pavement distress deterioration prediction
	Lopez et al. 2018, [28]	MGGP	Driving score calculation

* Oriented Genetic Programming
** Nested Align Genetic Programming

REFERENCES

[1] Koza, J. R., & Koza, J. R. (1992). Genetic programming: on the programming of computers by means of natural selection (Vol. 1). MIT press.

[2] Hrnjica, B., & Danandeh Mehr, A. (Eds.). (2018). Optimized Genetic Programming Applications: Emerging Research and Opportunities: Emerging Research and Opportunities. IGI Global.
[3] Mehr, A. D., Nourani, V., Hrnjica, B., & Molajou, A. (2017). A binary genetic programing model for teleconnection identification between global sea surface temperature and local maximum monthly rainfall events. Journal of Hydrology, 555, 397-406. [4] Mehr, A. D., Jabarnejad, M., & Nourani, V. (2019). Pareto-optimal MPSA-MGGP: A new gene-annealing model for monthly rainfall forecasting. Journal of Hydrology, 571, 406-415.
[5] Ahvanooey, M. T., Li, Q., Wu, M., & Wang, S. (2019). A Survey of Genetic Programming and Its Applications. TIIS, 13(4), 1765-1794.
[6] Mehr, A. D., & Kahya, E. (2017). A Pareto-optimal moving average multigene genetic programming model for daily streamflow prediction. Journal of hydrology, 549, 603-615.
[7] Ravansalar, M., Rajaee, T., & Kisi, O. (2017). Wavelet-linear genetic programming: A new approach for modeling monthly streamflow. Journal of hydrology, 549, 461-475.
[8] Heřmanovský, M., Havlíček, V., Hanel, M., & Pech, P. (2017). Regionalization of runoff models derived by genetic programming. Journal of hydrology, 547, 544-556.
[9] Chadalawada, J., Havlicek, V., & Babovic, V. (2017). A genetic programming approach to system identification of rainfall-runoff models. Water Resources Management, 31(12), 3975-3992.
[10] Ghorbani, M. A., Khatibi, R., Mehr, A. D., & Asadi, H. (2018). Chaos-based multigene genetic programming: A new hybrid strategy for river flow forecasting. Journal of hydrology, 562, 455-467.
[11] Mehdizadeh, S., & Sales, A. K. (2018). A comparative study of autoregressive, autoregressive moving average, gene expression programming and Bayesian networks for estimating monthly streamflow. Water resources management, 32(9), 3001-3022.
[12] Mehr, A. D. (2018). An improved gene expression programming model for streamflow forecasting in intermittent streams. Journal of hydrology, 563, 669-678.
[13] Hadi, S. J., & Tombul, M. (2018). Monthly streamflow forecasting using continuous wavelet and multi-gene genetic programming combination. Journal of hydrology, 561, 674-687.
[14] Jamei, M., & Ahmadianfar, I. (2019). Prediction of scour depth at piers with debris accumulation effects using linear genetic programming. Marine Georesources & Geotechnology, 1-12.
[15] Abdollahzadeh, G., Jahani, E., & Kashir, Z. (2017). Genetic programming based formulation to predict compressive strength of high strength concrete. Civil Engineering Infrastructures Journal, 50(2), 207-219.
[16] Velay-Lizancos, M., Perez-Ordoñez, J. L., Martinez-Lage, I., & Vazquez-Burgo, P. (2017). Analytical and genetic programming model of compressive strength of eco concretes by NDT according to curing temperature. Construction and Building Materials, 144, 195-206.
[17] Vanneschi, L., Castelli, M., Scott, K., & Popovič, A. (2018). Accurate High performance concrete prediction with an alignment-based genetic programming system. International Journal of Concrete Structures and Materials, 12(1), 72.
[18] Hodhod, O. A., Said, T. E., & Ataya, A. M. (2018). Prediction of creep in concrete using genetic programming hybridized with ANN. Computers and Concrete, 21(5), 513-523.
[19] Asim, K. M., Idris, A., Iqbal, T., & Martínez-Álvarez, F. (2018). Seismic indicators based earthquake predictor system using Genetic Programming and AdaBoost classification. Soil Dynamics and Earthquake Engineering, 111, 1-7.
[20] Qu, Z., Chen, Y. X., Liu, L., Xie, Y., & Zhou, Q. (2019). The Algorithm of Concrete Surface Crack Detection Based on the Genetic Programming and Percolation Model. IEEE Access, 7, 57592-57603.
[21] Goharzay, M., Noorzad, A., Ardakani, A. M., & Jalal, M. (2017). A worldwide SPT-based soil liquefaction triggering analysis utilizing gene expression programming and Bayesian probabilistic method. Journal of Rock Mechanics and Geotechnical Engineering, 9(4), 683-693.

[22] Soleimani, S., Rajaei, S., Jiao, P., Sabz, A., & Soheilinia, S. (2018). New prediction models for unconfined compressive strength of geopolymer stabilized soil using multi-gen genetic programming. Measurement, 113, 99-107.

[23] Kurugodu, H. V., Bordoloi, S., Hong, Y., Garg, A., Garg, A., Sreedeep, S., & Gandomi, A. H. (2018). Genetic programming for soil-fiber composite assessment. Advances in Engineering Software, 122, 50-61.

[24] Puente, C., Olague, G., Trabucchi, M., Arjona-Villicaña, P. D., & Soubervielle-Montalvo, C. (2019). Synthesis of Vegetation Indices using genetic programming for soil erosion estimation. Remote Sensing, 11(2), 156.

[25] Ranasinghe, R. A. T. M., Jaksa, M. B., Nejad, F. P., & Kuo, Y. L. (2019). Genetic programming for predictions of effectiveness of rolling dynamic compaction with dynamic cone penetrometer test results. Journal of Rock Mechanics and Geotechnical Engineering.

[26] Sahani, R., Ojha, A., & Bhuyan, P. K. (2017). Service levels of sidewalks for pedestrians under mixed traffic environment using genetic programming clustering. KSCE Journal of Civil Engineering, 21(7), 2879-2887.

[27] Chopra, T., Parida, M., Kwatra, N., & Chopra, P. (2018). Development of Pavement Distress Deterioration Prediction Models for Urban Road Network Using Genetic Programming. Advances in Civil Engineering, 2018.

[28] Lopez, J. R., Gonzalez, L. C., Wahlström, J., y Gómez, M. M., Trujillo, L., & Ramírez-Alonso, G. (2018). A Genetic Programming Approach for Driving Score Calculation in the Context of Intelligent Transportation Systems. IEEE Sensors Journal, 18(17), 7183-7192.