Engineering Applications of Artificial Intelligence 34 (2014) 66–74

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Prediction of slump flow of high-performance concrete via parallel hyper-cubic gene-expression programming

Li Chen [a,*], Chang-Huan Kou [b,1], Shih-Wei Ma [a,2]

[a] Department of Civil Engineering, Chung Hua University, No. 707, Sec.2, WuFu Rd., Hsinchu, Taiwan 30012, ROC
[b] Department of Architecture and Urban Planning, Chung Hua University, No. 707, Sec.2, WuFu Rd., Hsinchu, Taiwan 30012, ROC

ARTICLE INFO

Article history:
Received 4 September 2013
Received in revised form
25 March 2014
Accepted 6 May 2014
Available online 4 June 2014

Keywords:
Gene-expression programming
Slump flow
High-performance concrete
Optimization techniques
Nonlinear system
Parallel subpopulations

ABSTRACT

Parallel hyper-cubic gene expression programming (GEP) was constructed to estimate the slump flow of high-performance concrete (HPC) by using seven concrete ingredients. HPC is a highly complex material; because modeling its behavior is extremely difficult, robust optimization techniques are required. Because of complications caused by high dimensionality, obtaining globally optimal or nearly optimal solutions to such problems is extremely difficult. GEP is a type of evolutionary algorithm that simultaneously optimizes functions and their associated coefficients and is suitable for automatically discovering relationships between nonlinear systems. However, basic GEP, which generally suffers from premature convergence, is often trapped in local optima. Thus, in this study, various parallel subpopulations were processed to enhance search diversity and avoid local optima during GEP optimization procedures. The hyper-cubic topology rapidly spreads excellent solutions throughout all subpopulations. In addition, a migration mechanism, which exchanges chromosomes among the subpopulations, exchanges information during joint optimization to maintain diversity. In a case study, it was observed that for estimating HPC slump flow, the parallel hyper-cubic GEP is more accurate than the basic GEP and two types of regression model. Although both back-propagation neural networks (BPNN) and the proposed methods performed similarly, the proposed methods were preferable because formulas with measurable parameters were clearly provided.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Evolutionary computation techniques are based on the survival of the fittest, which is a powerful principle of evolution, and are very efficient optimization methods. Researchers successfully applied evolutionary algorithms to automatically generate programs or equations among the inputs and outputs. Since first proposed by Koza (1992), genetic programming (GP) has earned significant attention in terms of its ability to model nonlinear relationships. The focal point of the present research is to employ gene-expression programming (GEP) for the explicit mathematical formulation of the evapotranspiration. Gene-expression programming (GEP), a tree-based variant of genetic programming (GP), was first introduced by Ferreira (2001a). The basic difference between GEP and GP is due to the nature of the individuals: in GP the individuals are nonlinear entities of different sizes and shapes (parse trees); and in GEP the individuals are encoded as linear strings of fixed length (Ferreira 2001a, 2001b). GEP belongs to the class of evolutionary algorithms and, as such, shares the same parallelization features of evolutionary techniques. The parallel hyper-cubic topology of GEP herein presented is classified as a coarse or distributed and fine or cellular GEP, which considers several independent populations evolving in parallel with information exchange at predefined time slots.

High performance concrete (HPC) is a highly complex material manufactured of four to ten different components and its behavior is thus difficult to model. Traditional approaches model performance of concrete assuming an analytical equation followed by a regression analysis using experimental data to determine unknown coefficients in the equation (Mansour, 2004). However, the prediction ability of regression analyses may be limited due to highly non-linear problems (Baykasoğlu et al., 2004). In recent years, artificial neural networks (ANNs) performed exceptionally as regression tools because they are more direct than the traditional statistical methods (Yeh, 2006a). Moreover, ANNs are highly nonlinear and can capture complex interactions between input/output variables in a system, without prior knowledge about the nature of interactions. The main advantage is that ANNs do not

* Corresponding author. Tel.: +886 3 5186718.
E-mail addresses: lichen@chu.edu.tw (L. Chen), chkou@chu.edu.tw (C.-H. Kou), d09404006@chu.edu.tw (S.-W. Ma).
[1] Tel.: +886 3 5186706.
[2] Tel.: +886 3 5186718.

http://dx.doi.org/10.1016/j.engappai.2014.05.005
0952-1976/© 2014 Elsevier Ltd. All rights reserved.

explicitly assume a model form, which is a prerequisite in the parametric approach (Ji et al., 2006). However, such "black box" models do not easily generate analytical equation forms (Baykasoğlu et al., 2004). Chen (2003) used GP to model the strength of HPCs and demonstrated that the GP method obtains high estimation accuracies. Several other studies including Tsai and Lin (2011) have employed improved GP to modeling concrete materials. Yeh and Lien (2009) and Peng et al. (2010) proposed using genetic operation trees (GOT) to study concrete strength.

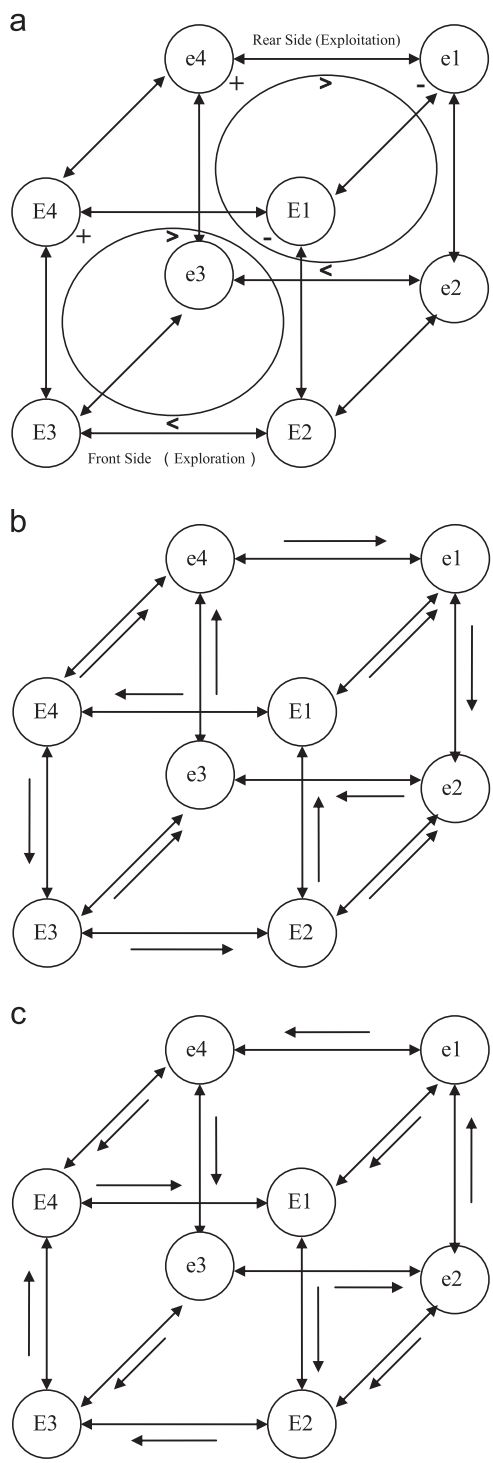

Fig. 1. Structure of a hyper-cubic GEP: (a) basic topology; (b) refinement; (c) expansion.

Baykasoğlu et al. (2009) employed GEP (first proposed by Ferreira (2001a, 2001b)) for concrete strength, cost and slump. Six input variables were used in high-strength concrete (HSC) mixture: water (*W*), cement (*C*), fine aggregate (s/a), coarse aggregate (c/a), fly ash (FA), air entraining agent (AE), and superplasticizer (SP). There were 104 mix designs to investigate the performance of the GEP algorithm. The population size was set as 100 and ran with generations between 3000 and 20,000 and the mathematical operators included {+, −,*, /, √,e, log, sin, cos, tan}. The best results obtained from the GEP tests were 4, 427.44, 0.04 MSE for compressive strength, slump, and cost respectively. It shows that the result of slump model is the worst one which means it is very difficult to predict the values of slump or slump flow of concrete.

The main purpose of this paper is to improve the modeling techniques of estimating HPC parameters by a parallel structure of GEP. Because of the complex nonlinear relationship between several ingredients and the slump flow of HPC, basic GEP usually suffers from premature convergence which cannot acquire satisfying solutions. The current study addressed this problem by enhancing the ability of the algorithm to escape from local optimal solutions. In the next section, we begin by introducing the implementation of the parallel hyper-cubic structure of GEP to increase the diversity of solutions for obtaining the optimal equation. The section after that demonstrates a case study of the slump flow model of HPC. The results of the parallel hyper-cubic GEP are compared with those of basic GEP, two types of multiple linear regression (MLR) and back-propagation neural networks (BPNN). Finally, the present study provides conclusions and closing remarks.

2. Parallel hyper-cubic gene-expression programming (GEP)

2.1. Basic GEP

Gene expression programming (GEP) (Ferreira, 2001b) uses linear chromosomes. A chromosome is composed of genes containing terminal and nonterminal symbols. Establishing a symbolic model using GEP needs to determine appropriate mathematical operations, variables and constants. When the linear string is set up to represent a specific mathematical formula, it can generate predicted output value for each data.

2.2. Parallel structure

The search process would be ineffective iterations of GEP operators within the vast chromosomes and result in very slow evolution between consecutive generations, consequently sharply decreasing the possibility of converging to an optimal solution. Therefore, GEP is likely to be trapped in a region that does not contain the global optimum. This problem, called premature convergence, has been recognized as a serious failure mode for genetic-based algorithms (Eshelman and Schaffer, 1991). A multi-population GA divides a single population into smaller subpopulations. Each subpopulation evolves by genetic operations in parallel with the other, while maintaining a limited but powerful interaction between all subpopulations. The most important advantage of subpopulations is the enhanced diversity among the subpopulations (Herrera and Lozano, 2000). The similar parallel structure can be used to GEP by preserving diversity due to the semi-isolation of the subpopulations. The parallel implementation was implemented on a PC four-core machine with 16 G bytes of memory. Therefore, parallel GEP were adopted to improve the search diversities and described as follows.

Download English Version:

https://daneshyari.com/en/article/380430

Download Persian Version:

https://daneshyari.com/article/380430

Daneshyari.com