Proceeding Series of the Brazilian Society of Applied and Computational Mathematics, v. 6, n. 1, 2018.

Trabalho apresentado no XXXVII CNMAC, S.J. dos Campos - SP, 2017.

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

A New Approach To Software Effort Estimation Using Linear Genetic Programming

Roberto Sánchez[1]
Diego P. Pinto-Roa[2]
Facultad Politécnica
Universidad Nacional de Asunción - UNA, San Lorenzo, Paraguay

1 Resumo

Accurate estimation of effort in software projects is one of the most desired capabilities in software development organizations and is a big challenge for projects manager. Effort is measured in terms of person months and is evaluated with the help of different methods. COCOMO is the most popular non-proprietary software estimation model in literature as well as in industry [1] [2]. Both underestimated and overestimated effort are harmful for projects and caused development schedule delay and poor software quality. There are many software cost estimation models have been developed over the last decades. The computing intelligence techniques have been grown. Some researches showed a promising research direction in software cost estimation with genetic programming [4]. Linear Genetic Programming (LGP) is a genetic programming variant that evolves sequences of instructions from an imperative programming language. This paper presents a LGP approach to evaluates and compares the potential on software project data set. The objective in this paper is the finding of new models from historical data using LPG that evolves sequences of instructions from an imperative programming language, like C, C++, Java, etc [3]. In this work, programs are represented as sequences of instructions that accept a minimum number of variables, called registers. The instruction form may be the following: $r_0 = r_1 + r_5$ where r_i is a register. This type of representation includes an operator on operand registers. In this work implementation, each instruction of a program includes four values (an operation, a destination and two operands). For instance, a 4-tuple as $< +, 0, 2, 9 >$ represents the instruction $r[0] = r[2] + r[9]$. New populations are produced by selecting with larger probability good ranking solutions and performing genetic operations to produce offspring solutions that let the population evolve. This process is repeated over many generations until some stop condition is reached. Experiments in this paper have been conducted from COCOMO database [1], a data set publicly available which consists of 63 projects. Data were divided into training and validation periods. The objective

[1]totosanchez@gmail.com
[2]dpinto@pol.una.py

© 2018 SBMAC

function measures the prediction quality of an individual or program. A popular error function for estimation problems is the mean magnitude relative error (MMRE). In the proposed model, MMRE is set as the Objective Function to minimize. The LGP program is implemented in C#.NET 2010 programming environment. The results of experiment in Table 1 show that the proposed model holds lower MMRE than COCOMO model with accuracy of MMRE between 1% and 30% of fifteen estimation models. Results highlight LGP potential as a method to produce models for software effort estimation. As future work, the authors planned to continue testing and analyzing other data sets. Also, they intend to compare LPG to other techniques such as Neural Networks or similar.

Table 1: Estimation Results

Model Used	$\sum MMRE$	$MMRE$	% Improvement
Cocomo	4,592289781	0,38269082	-
LGP 1	3,202167159	0,26684726	30,27
LGP 2	4,326669957	0,36055583	5,78
LGP 3	2,994904485	0,24957537	34,78
LGP 4	4,34638703	0,36219892	5,35
LGP 5	3,282077133	0,27350643	28,53
LGP 6	4,442038579	0,37016988	3,27
LGP 7	4,362001784	0,36350015	5,01
LGP 8	4,218742431	0,35156187	8,13
LGP 9	3,436702829	0,2863919	25,16
LGP 10	4,540089824	0,37834082	1,14
LGP 11	4,325920876	0,36049341	5,80
LGP 12	4,213055967	0,351088	8,26
LGP 13	3,292910739	0,27440923	28,29
LGP 14	4,10875525	0,34239627	10,53
LGP 15	4,508219057	0,37568492	1,83

References

[1] B. Boehm. Software Engineering Economics*Englewood Cliffs*, NJ. Prentice-Hall, 1981.

[2] B. Boehm, R. Valerdi. Achievements and Challenges in Cocomo-Based Software Resource Estimation*IEEE Software*, pp. 74-83, 2008.

[3] M. Brameier and W. Banzhaf, Linear Genetic Programming (Genetic and Evolutionary Computation), *Springer*, EEUU, 2014.

[4] A. F. Sheta. Estimation of the cocomo model parameters using genetic algorithms for nasa software projects, *Journal of Computer Science*, volume 2, no. 2, pp. 118-123, 2006.

© 2018 SBMAC