A Generic Analogy-Centered Software Cost Estimation Based on Differential Evolution Exploration Process

ZAHID HUSSAIN WANI^{D*}, JAVAID IQBAL BHAT AND KAISAR JAVEED GIRI

Department of Computer Science, Islamic University of Science and Technology, Awantipora, Jammu and Kashmir 192122, India *Corresponding author: zahid.uok@gmail.com

Software cost estimation is the prediction of development effort and calendar time required to develop any software project. It is considered to be the very fundamental task for successful execution of an on-going project as well as budgetary requirements of futuristic projects. As accuracy in software cost estimation is very hard because of the availability of vague information at the time of inception of the software project, it prompted many researchers to explore in this domain from past decades. Their pioneer works suggest a bulk of techniques for this purpose. However, because of the availability of large number of estimation techniques, it becomes hard for any software practitioners to select an appropriate one. To help the industry practitioners in these situations, a novel analogy-centered model based on differential evolution exploration process is proposed in this research study. The proposed model has been assessed on 676 projects from 5 different data sets and the results achieved are significantly better when compared with other benchmark analogy-based estimation studies. Furthermore, being the very less computational cost of the proposed model, it is suggested that the proposed model be considered as the preliminary stage of any analogy-based software estimation technique.

Keywords: analogy-based estimation; differential evolution; software cost estimation; genetic algorithms

Received 31 May 2020; Revised 16 September 2020; Accepted 29 November 2020 Handling editor: Sabah Mohammed

1. INTRODUCTION

Software cost estimation is the prediction of both building effort and calendar time required to plan, design and develop a software project. The building effort includes the sum of hours, days or months of human work. From the very inception of software project development, software development organizations used to struggle against the challenges being forwarded by accurate estimations.

A good estimation model assists any software project manager in accurate estimation of both delivery time and development cost of a software development project as both being dependent on good estimation process. Conversely, an accurate estimate becomes the key thing used by software development project managers in evaluating project progress, thereby, yielding them a good track of potential cost control and delivery accuracy. Furthermore, an accurate estimate in broader perspective will give the organization a better understanding of resource utilization and noticeably will land the organizations in a better schedule of their futuristic projects. However, with mere selection of one wrong cost estimation factor, the cost estimation could result in an inaccurate estimate, which will affect the allotted funds for the project [1] and accordingly will result in wastage of resources and so in its timing constraint. As an illustration, the famous project namely incomplete Checkout Launch Control System of National Aeronautics and Space Administration (NASA) was canceled down soon after their initial \$200 million estimate when surpassed by another \$200 million [2].

In this paper, analogy-based cost estimation is considered as most of works based on analogy are already available [3, 4]. Further, in this research study, first estimations based on parametric equation have been totally ridiculed as they demand the data to be articulated strictly as per the parametric model and second thing that has been carried out in this analogybased estimation (ABE) study is that genetic algorithms used