Available online at www.sciencedirect.com

SciVerse ScienceDirect

International Journal of Project Management 30 (2012) 967 – 977

International Journal of Project Management

www.elsevier.com/locate/ijproman

Forecasting enterprise resource planning software effort using evolutionary support vector machine inference model

Jui-Sheng Chou [a,*], Min-Yuan Cheng [a], Yu-Wei Wu [a], Cheng-Chieh Wu [b]

[a] National Taiwan University of Science and Technology, Taipei, Taiwan
[b] Data Systems Consulting Co., Ltd., Taichung, Taiwan

Received 20 June 2011; received in revised form 24 January 2012; accepted 16 February 2012

Abstract

Despite significant advances in procedures that facilitate project management, the continued reliance of software managers on guesswork and subjective judgment causes frequent project time overruns. This study uses an Evolutionary Support Vector Machine Inference Model (ESIM) for efficiently and accurately estimating the person-hour of ERP system development projects. The proposed ESIM is a hybrid intelligence model integrating a support vector machine (SVM) with a fast messy genetic algorithm (fmGA). The SVM mainly provides learning and curve fitting while the fmGA minimizes errors. The analytical results in this study confirm that, compared to artificial neural networks and SVM, the proposed ESIM provides preliminary prediction at early phase of ERP software development effort for the manufacturing firms with superior accuracy, shorter training time and less overfitting. Future research can develop user-friendly expert systems with window or browser interfaces that can be used by planning personnel to flexibly input related variables and to estimate development effort and corresponding project time/cost.

Keywords: Enterprise resource planning; Software effort prediction; Project management; Hybrid intelligence

1. Introduction

Enterprise resource planning (ERP) enables developers to enhance the global competitiveness and sustainability of their client enterprises by ensuring efficient resource allocation. In practice, ERP software functions and specifications are highly unpredictable at early stages of R&D; thus, initial cost estimation relies mostly on the subjective judgments of experienced software engineers. Although knowledgeable sales managers or estimators may generate accurate cost assessments *via* a cooperative approach, professionals in small and medium-sized software enterprises are often difficult to train and highly mobile. Thus, the difficulty of retaining experienced personnel with project knowledge results in such problems as loss of project know-how.

Despite the significant advances in the procedures that facilitate project management (PMI, 2008), product managers in the software industry still encounter problems requiring guesswork and subjective judgment, which often result in inaccurate estimates. Effort estimation is not functionally related to the basic drivers of ERP system development. Companies can thus lose market share and orders when attempting to attract customers during the early marketing phase. Although human experts can achieve satisfactory outcomes, shortfalls typically result from inefficient information management. Shortcomings in current subjective assessments or analogous methods indicate the urgent need and opportunity for improvement.

Unlike traditional manufacturing, most software product development costs are incurred by investment in human resources. As software is a virtual intelligence and customer service–oriented product, software developers must estimate project completion time at early stages. Many studies (Ahmed and Muzaffar, 2009; de Barcelos Tronto et al., 2008; Elish, 2009; Finnie et al., 1997; Huang and Chiu, 2006; Huang et al., 2008; Jørgensen, 2010; Kazemifard et al., 2011; Lopez-Martin, 2011; Mair et al., 2000; Oliveira et al., 2010; van Koten and Gray, 2006) have proposed cost or effort forecasting methods for software development

* Corresponding author at: 43 Sec. 4, Keelung Rd., Taipei 106, Taiwan. Tel.: +886 2 2737 6321; fax: +886 2 2737 6606.
E-mail address: jschou@mail.ntust.edu.tw (J.-S. Chou).

0263-7863/$36.00
doi:10.1016/j.ijproman.2012.02.003

projects. Nevertheless, the effectiveness and efficiency of approximate inference methods for estimating ERP system development efforts *via* hybrid intelligence (*i.e.*, combination of multiple artificial intelligence techniques) are rarely addressed.

Developing deterministic mathematical models for solving project prediction problems is both difficult and expensive. Approximate inference, a fast and cost effective approach, is a viable alternative to deterministic mathematical modeling. Inference is the process of deriving new knowledge from known information. As the known information changes, the inference process adapts accordingly. Prediction problems are complex and often involve substantial uncertainty, vagueness, and incomplete or inexact data. Therefore, the inference process must fit environmental conditions (Mareels and Polderman, 1996). Humans can process and solve complex problems, even those involving uncertainty, imprecision, and incomplete information. Therefore, imitating the process of human inference is an effective approach to solving project prediction problems.

The primary objective of this research was predicting development time for ERP software project by using an Evolutionary Support Vector Machine Inference Model (ESIM). After reviewing the relevant literature and collecting data and ERP software specifications from past projects in the manufacturing industry, an ESIM was proposed for simultaneously searching for the fittest support vector machine (SVM) parameters within a globally optimized model. An early accurate time estimate during the negotiation stage with potential clients can eliminate unnecessary bargaining and cost changes in subsequent processes.

The rest of this paper is organized as follows. Section 2 reviews pertinent literature in software project estimation. Section 3 then discusses the research methodology and model adaptation process. Next, Section 4 illustrates the applicability of the ESIM technique in a case study of a leading ERP software system provider. Conclusions are finally drawn in Section 5 along with managerial implications and suggestions for future research.

2. Literature review

The ERP software systems typically manage internal enterprise resources and departmental business operations. Functions include business management, production management, financial planning, quality management, and human resource management. The typical phases of a software development project (Yeo, 2001) are, in chronological order, system analysis, system design, programming, testing, deployment, and online operations.

Estimation accuracy heavily depends on the amount and quality of information available at estimation time. Thus, providing accurate preliminary project estimates is extremely challenging, particular during the software development phases (Lopez-Martin, 2011). A major limitation of current practices is that, when contacting clients for make-to-order quotes, project managers must estimate time-to-complete based solely on known product attributes whereas virtually all software costs at early stages of software development are implicitly determined by investment in human resources.

High-quality ERP software development projects require efficient project control mechanisms to schedule and manage software development progress. Project estimation is thus a prerequisite for all subsequent planning activities. Although none of the studies have quantified the effort or duration requirements for ERP software development, project estimation schemes proposed for other software project and industries include case-based reasoning (Belecheanu et al., 2003; Chou, 2009b; Finnie et al., 1997; Mendes et al., 2002, 2003; Yang and Wang, 2008), multiple regression analysis (Adriano L.I, 2006; Bashir and Thomson, 2001, 2004; M. Camargo et al., 2003; Caputo and Pelagagge, 2008; Chou, 2009a; Chou et al., 2006; Ferens, 1998; Finnie et al., 1997; Huang et al., 2008; Marban et al., 2008; Mittas and Angelis, 2010; Sentas and Angelis, 2006), activity-based costing (Baykasoglu and Kaplanoglu, 2008; Ben-Arieh and Qian, 2003; Tornberg et al., 2002), artificial neural networks (Attarzadeh and Ow, 2010; Berlin et al., 2009; Caputo and Pelagagge, 2008; de Barcelos Tronto et al., 2008; Delen et al., 2005; Finnie et al., 1997; Huawang and Wanqing, 2008), support vector machine (Adriano L.I, 2006; An et al., 2007; Lu and Tsai, 2008; Pal and Deswal, 2008), genetic algorithms (Huang and Chiu, 2006; Oliveira et al., 2010) and function point analysis (Albrecht and Gaffney, 1983; Finnie et al., 1997; IFPUG, 2009; Longstreet, 2002; Mendes et al., 2003; Myrtveit et al., 2005; Yinhuan et al., 2009). However, the above works considered single techniques rather than hybrid intelligence schemes.

Artificial intelligence (AI)-based approaches are related to computer system designs that attempt to resolve problems intelligently by emulating human brain processes. As AI technology enhances the ability of computer programs to handle tasks for which humans are still superior (Haykin, 1999), AI models are typically used to solve project estimation problems. Various scientific and engineering fields have recently combined different AI paradigms to enhance efficacy. Numerous studies confirm that hybrid AI schemes outperform single techniques in project estimation (Chen, 2007; Kim and Shin, 2007; Lee, 2009; Li et al., 2005; Min et al., 2006; Nandi et al., 2004; Wu, 2010; Wu et al., 2009). Fast messy genetic algorithm (fmGA) (Goldberg et al., 1993) and the SVM (Vapnik, 1995) are two such tools that have proven effective for solving various project management problems.

Specifically, the SVM in the ESIM is used mainly for learning and curve fitting while the fmGA optimizes prediction error. Given the characteristics and merits of fmGA and SVM, the ESIM proposed in this study combines them for predicting development effort in ERP software projects. Fig. 1 shows the software project quotation and engineer-to-order processes integrated with the proposed ESIM. This ESIM was designed to achieve the fittest C and gamma parameters with minimal prediction error. The proposed approach considers implicit knowledge from historical cases in the software industry so that project managers or decision-makers can overcome challenges during early cost quotation.

3. Evolutionary support vector machine inference model

3.1. Support vector machine integrated with fast messy genetic algorithm

Support vector machine was first introduced by Vapnik (1995) and colleagues at AT&T Bell Laboratories (Vapnik,