Expert Systems with Applications 36 (2009) 8938–8945

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Recommender system for software project planning one application of revised CBR algorithm

Heng-Li Yang [a,*], Cheng-Su Wang [b]

[a] Department of Management Information Systems, National Cheng-Chi University, 64 Section 2, Chihnan Road, Mucha Dist., Taipei 116, Taiwan
[b] Department of Information Management, Chung Hua University, 707, Sec. 2, WuFu Rd., Hsinchu, Taiwan

ARTICLE INFO

Keywords:
Project planning
Case-based reasoning
Hybrid AI applications
Recommender system
Multiple objective decisions

ABSTRACT

Project management is an experience-driven and knowledge-centralized activity. Therefore, project managers require some assistance to reduce the uncertainty at the early stage of constructing project plans. To overcome the predicament faced by project managers, this investigation proposes a hierarchical criteria architecture (HCA) to enable project managers to describe project requirements adequately. Furthermore, to solve HCA problems, a revised case-based reasoning (RCBR) algorithm, is presented and a recommender system for software project planning is implemented, based on multiple objectives decision techniques and the mining approach. Finally, the proposed RCBR algorithm is successfully applied to analyze 41 real projects from a software consultancy in Taiwan. Experimental results demonstrate that RCBR can efficiently provide related information to help project managers to construct project plans at an early stage. Additionally, the knowledge discovery process of RCBR provides project managers with results similar to what-if analysis. The knowledge can enable project managers to obtain feasible information to re-schedule project resources, and bargain with their customers in the early project planning stage.

1. Introduction and related work

Information technology (IT) or information system (IS) project management is regarded as a knowledge-centralized and experience-driven activity that is hard for inexperienced staff to implement (Lee & Lee, 2006). The Standish Group CHAOS report indicated that only 47% of software project could deliver within budget and required functions (Standish Group International, 2004). Additionally, Tesch, Kloppenbborg, and Frolick (2007) even found that only 28% of IT projects are successfully completed on time. IT project management is a difficult and challenging job, often with a low success rate. Therefore, project managers (PMs) of IT/IS projects indeed require some assistance to do their jobs well (Dweiri & Kablan, 2006). However, the entire IT project management process is highly uncertain and full of risks, and success rate of a project generally depends on various critical factors. A systematic methodology is required to solve problems faced by PMs.

Some researchers have, in recent decades, addressed project management issues from different perspectives. Some researchers have focused on improving project information transparency and project quality to enable PMs to increase the project success rate. Raymond and Bergeron (2008) found that project management information system (PMIS) directly influences project success by complying with project deadlines and technical specifications. The PM certainly requires additional assistance even when not applying PMIS (Raymond & Bergeron, 2008). Garcia, Roman, Penalvo, and Bonilla (2008) addressed this problem using association mining rules that determine meaningful project strategies from discrete numerical project data. Additionally, Stamelos and Angelis (2001) successfully applied analogy-based methodology for cost estimation across multiple software projects by project portfolios. The above researchers intended to improve information transparency and quality via prediction (such as budget or project duration prediction) and IT systems (such as PMIS or knowledge sharing mechanism). Conversely, many researchers have focused on project risk analysis, which has become the primary research stream within project management. Wallace, Keil, and Rai (2004) and Han and Huang (2007) derived software project risk factors by cluster analysis and empirical surveys, respectively. Both these two researches found that requirement and project planning and control are closely related to project requirement risk and complexity risk level (from low to high). Moreover, Rodriguez-Repiso, Setchi, and Salmeron (2007) considered these factors as critical success factors, and adopted fuzzy cognitive map to build an IT project success model. All the cited investigations above involve predicting the level of project risk, and then evaluating the project success rate.

* Corresponding author. Tel.: +886 2 29387651; fax: +886 2 29393754.
E-mail addresses: yanh@nccu.edu.tw (H.-L. Yang), cswang@chu.edu.tw (C.-S. Wang).

0957-4174/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2008.11.050

A software project management has four stages, namely project planning, scheduling, monitoring and control. To evaluate the project success, both the project risk factor and project information quality issue should be addressed in all four stages. Each stage, particularly project planning, is closely related to project success (Lee & Lee, 2006). A project without adequate project planning has a high risk level (Han & Huang, 2007; Wallace et al., 2004). Thus, a comprehensive project planning could lead to project success. However, project planning should consider different project attributes, such as type, size and customer requirements. The situation of project planning is highly uncertain and vague during the early stage. Therefore, PMs need a lot of assistance and information.

Tina, Ma, and Liu (2002) integrated knowledge rules into a mathematical model for R&D to assist project selection for Nation Science Foundation of China. However, mathematical models typically make too many assumptions to be applied well. Since PMs frequently construct project plans based on past experiences, Lee and Lee (2006) recommended adopting case-based reasoning (CBR) algorithm to provide previous project cases requiring the least modification for constructing new project plans. CBR appears to be a feasible algorithm to solve the project planning problem with multiple conflicting objectives. Unfortunately, project requirements of real applications are too complicated to describe via multiple objectives. A sophisticated model is required to permit PMs to present their project requirements adequately. Additionally, PMs applying the proposed model of Lee and Lee (2006) might have no idea how to revise the retrieved case, even if it is the fittest one. Yang, Yin, Lin, and Pan (2007) claimed that only the actionable information would provide reference value for decision makers. PMs should obtain refined information, rather than raw cases, from the case base. Therefore, this study develops a systematic methodology to assist project planning.

PMs should be aware of all attributes of projects, such as project size (Garcia, Quintales, Penalvo, & Martin, 2004; Martin, Pearson, & Furumo, 2007), cost (Love & Irani, 2003), budget, customer requirements (Dweiri & Kablan, 2006) and customer profiles (Dweiri & Kablan, 2006) when constructing project plans. PMs should ideally perform what-if analysis to build the project plan prototypes, such as recovery plans and staff schedules. These activities are experience-driven and have uncertain results, and become increasingly complicated and difficult due to globalization (Mahaney & Lederer, 2003). For instance, resources (such as staff, information and knowledge) might need to be transferred among countries and time zones. Therefore, early planning of IT projects is increasingly important.

Artificial intelligence (AI) approaches can effectively reduce the uncertainty, thus enabling PMs to handle the early stage of project planning comfortably. Donzellli (2006) presented a decision support system (DSS) to assist PMs to predict and simulate project implementation processes. Donzellli's DSS successfully improved the efficiency of project management and the schedule control quality. Dweiri and Kablan (2006) also applied fuzzy decision-making techniques to project management. They proposed a fuzzy system to evaluate and predict the efficiency of project management through project cost, project time and project quality.

This study introduces an AI approach, called case-based reasoning (CBR), to decrease uncertainty in early stages of project planning. As is well-known, CBR is particularly useful for solving ill-defined and unstructured problems (Belecheanu, Pawar, Barson, Bredehorst, & Weber, 2003). CBR is the most appropriate AI technique for producing project planning recommendations at early stages (Lee & Lee, 2006), since IT (or IS) projects are typically described in an unstructured manner, such as in the case-format. CBR is a paradigm, concept and instinctive mechanism for problem solving. Similar to the human problem solving process, CBR retrieves past experiences for reuse on a target problem (Yang & Wang, 2008). Cirovic and Cekic (2002) adopted CBR to support the preliminary design phase of a construction project. They retrieved previous cases from a historical integrated project database, then stored them into a knowledge base for quality designing. The retrieved solutions from past cases may need to be revised for further application. The successful problem-solving experiences are then retained for further reuse. Traditional CBR algorithms have 4R-reasoning steps (Aamodt & Plaza, 1994). However, although the project recommender system may provide the relevant case(s) for PMs referencing, the revised directions of these retrieved IT projects remain ambiguous. PMs still do not know how to adjust these projects to generate a reasonable project plan prototype.

This study presents a revised CBR mechanism, named RCBR, that integrates the traditional CBR algorithm with mining approaches. RCBR consists of two stages. In stage I, RCBR retrieves the fittest case(s) for planning reference according to the query requirement of PMs. In stage II, RCBR mines the cases retrieved in stage I to provide further refined and potential knowledge. The recommendation result of RCBR could decrease the uncertainty in early planning stages, and thus assist PMs to plan projects successfully. Additionally, this investigation has the following goals:

(1) to enable PMs to consider their project requirements carefully, so that the retrieved case(s) fit reality;
(2) to revise the traditional CBR approach, and apply RCBR to reduce uncertainty in project planning, and
(3) to provide some clues (or information) for revised direction, thus improving the usefulness of the online application.

The remainder of this study is organized as follows. Section 2 presents a new problem description called hierarchical criteria architecture (HCA). Section 3 describes the proposed revised-CBR algorithm. Section 4 summarizes the experimental results of the proposed model. Section 5 presents conclusions and recommendations for future research.

2. Suggested problem presentation: hierarchical criteria architecture

The problem descriptions in traditional CBR algorithms have to follow the structure and stored format of case descriptions. In other words, PMs have to raise their queries according to the case structure, regardless of their real requirements. For instance, Fig. 1 depicts a traditional CBR, CBR-Works, which formulates PMs' queries in a multi-objective format (Coello, 2000). However, as the decision conditions become increasingly complex, the problem descriptions in one-level multiple objectives become too rough to represent the decision problem correctly and completely. The solution is difficult to trust if the decision problem cannot be described well.

Adomavicius and Tuzhilin (2005) stated that a next-generation recommender system should be able to solve multiple dimensional problems, which are complex and close to the real situations faced by decision makers. To provide PMs with the most appropriate case(s) for project planning references, a new problem presentation model is stipulated to permit PMs to focus adequately on their project requirements.

This study proposes a novel problem description, named hierarchical criteria architecture (HCA), to enable decision makers to describe their decision problems adequately. As implied by the name, HCA allows decision makers to describe each decision objective with multiple levels to their desired level of details. If a problem is described in HCA, then decision makers can drill each problem dimension down until the required level of detail is reached. Additionally, PMs could assign different relative weights to each attribute to represent its significance. Fig. 2 illustrates PMs evaluating the case similarity in terms of three dimensions, "customer",