

기술적 분석을 이용한 프로젝트 비용 산정 보정 기법

양희문^{0,a}, 유준범^b, 차성덕^a

^a고려대학교 컴퓨터정보통신대학원 소프트웨어공학과, ^b건국대학교 컴퓨터공학부
yanghm77@korea.ac.kr, jbyoo@konkuk.ac.kr, scha@korea.ac.kr

An Adjusting Method on Project Cost Estimation based on Technical Analysis

Heemoon Yang^{0,a}, Junbeom Yoo^b, Sungdeok Cha^a

^aDepartment of Software Engineering, Graduate School of Computer Information & Communication, Korea University.

^bDivision of Computer Science and Engineering, Konkuk University.

1. 서론

소프트웨어 프로젝트에서의 비용 보정은 이해당사자의 견해 차이와 시장 경쟁의 논리 같은 비기술적 요인들을 기반으로 진행되는 문제가 있다. 한국 소프트웨어 진흥원 정책연구[1]에서는 ‘적절치 않은 발주금액(53.7%)’을 불공정 거래의 가장 큰 문제로 분석하였다. 이는 Top-down, 계수 모델, Bottom-up와 같은 기법을 사용하여 비용 산정을 수행하지만, 이후 협상에서 비기술적인 비용 보정을 수행하기 때문이다.

본 연구에서는 기술적 분석을 이용한 비용 보정 기법을 제안하고, 유효성을 검증하였다. 기술적 분석은 프로젝트에 영향을 미치는 특성과 각각의 특성을 척도로 분류하는 것을 포함한다. 보정 기법은 기술적 분석을 기반으로 참조 모델과 예측 알고리즘을 포함한다. 참조 모델은 실적 데이터를 기반으로 특성들의 가중치와 분류한 척도 항목들의 평균을 도표로 나타낸 것이다. 예측 알고리즘은 참조 모델을 기반으로 신규프로젝트의 비용 보정 비율을 수식으로 나타낸 것이다. 제안한 보정 기법은 예측 데이터의 평균 오차율이 $\pm 4.32\%$ 로 유효성이 높음을 검증하였다.

2. 기술적 분석을 이용한 보정 기법

제안된 기법은 기술적 분석을 이용하여 비용을 보정한다. 기술적 분석은 프로젝트에 영향을 미치는 특성들을 정의하며, 각각의 특성들을 척도로 분류하는 것을 포함한다. 보정 기법은 기술적 분석을 기반으로 신규 프로젝트의 비용 보정 비율을 예측하는데 사용하며, 참조 모델과 예측 알고리즘을 포함한다.

특성은 프로젝트의 3가지 상태를 취합한 요소와 가중치를 정의한 것이다. CHAOS Report에서는 프로젝트 상태를 3가지로 분류하여 프로젝트의 성공률과 실패율을 조사하였다[2]. 프로젝트의 3가지 상태는 성공(최초 정의한 특징과 기능을 만족하면서 주어진 일정과 예산안에 완료되었을 경우), 도전(최초 명세 된 것보다 적은 특징과 기능을 제공하며 일정과 예산이 초과한 상태에서 완료되었을 경우), 실패(개발과정 중 특정 단계에서 취소되거나 중단되었을 경우)이다.

분류는 특성들을 척도로 나타낸 항목들이다. ‘요구사항 명확도’ 특성은 요구 공학 프로세스의 요소인 타당성 조사, 추출과 분석, 명세화, 검증으로 척도를 분류하였다. ‘고객 PM 참여도’ 특성은 프로젝트 수행 조직의 구조인 기능형 조직, 약/중/강 매트릭스 조직, 전담 조직으로 척도를 분류하였다. 나머지 특성은 태도나 가치를 수치화하는데 널리 쓰이는 Likert 5-Point 척도를 사용하여 분류하였다.

참조 모델은 실적 데이터를 기반으로 특성들의 가중치와 분류한 척도 항목들의 평균을 도표로 나타낸 것이다. 그림 1은 비용 보정에 사용될 참조 모델을 생성한 예제이다. 모델을 살펴보면, 과거에 수행된 프로젝트 A, B, C, D, E의 산정비용(보정 전)과 확정비용(보정 후)에 대한 비율(확정비용/산정비용)을 보여준다. 참조 모델의 데이터는 PM의 설문조사를 수행하여 데이터베이스화하였다. 데이터베이스를 토대로 분류 항목의 선택된 개수를 Count하고 유효수를 구하여 평균(SUM/유효수)을 구하였다. 가중치는 프로젝트의 특성 항목에 대한 값을 나타내고 있다.

예측 알고리즘은 참조 모델을 기반으로 신규프로젝트의 비용 보정 비율을 수식으로 나타낸 것이다. 신규 프로젝트가 발생하면 분류된 특성들의 척도 항목을 판단한다. 판단한 항목은 참조 모델에서 특성의 분류 항목에 해당하는 평균과 가중치를 곱하고 특성 수만큼 반복하여 더한다. 더한 값을 전체 가중치의 합으로 나누어 PV를 구한다. 그림 1의 예측 알고리즘은 PV(Prediction Variable)를 구하는 수식을 보여주고 있다.

PJT	산정가 (천원)	확정가 (천원)	비율 (%)	요구사항 명확도				고객PM 참여도				
				1.타당성 조사	2.추출과 분석	3.명세화	4.검증	1.기능 조직	2.약 매트릭스	3.중 매트릭스	4.강 매트릭스	5.전담 조직
A	1,000	1,000	1.00	1.00				1.00				
B	2,000	1,400	0.70				0.70					0.70
C	1,000	800	0.80			0.80				0.80		
D	3,000	2,700	0.90		0.90				0.90			
E	2,000	1,500	0.75			0.75						0.75
유효수				1	1	2	1	1	1	1	1	1
평균(SUM/유효수)				1.00	0.90	0.78	0.70	1.00	0.90	0.80	0.75	0.70
가중치				20.00				15.39				

$$PV = \sum_{i=1}^n A_i \cdot W_i / \sum_{i=1}^n W_i$$

n: Characteristic Count, A: Average, W: Weight

그림 1. 참조 모델과 예측 알고리즘

3. 실험 결과

제안한 보정 기법은 실험으로 유효성을 검증하였다. 실험은 2곳의 SI 시스템 개발 업체들에서 2005 ~ 2008년 수행한 46건의 실적 데이터를 사용하였다.

실험 데이터는 앙상블 학습인 Bagging[3]을 기반으로 독립적인 Training Data와 Test Data로 분리하였다. Bagging(Bootstrap+ Aggregating)은 예측 데이터를 분류기 수만큼 반복 선택하여 학습 데이터 집합을 만들고서 각각의 학습 데이터로 개별 분류기를 학습한다. 이렇게 학습 된 분류기의 결과를 합산하여 최종 결과를 얻는다. Bagging 을 기반으로 5번 반복하여 독립적인 Training Data Set과 Test Data Set을 추출하였다. 추출된 Test Data Set의 평균 비율은 35.65%로 Bagging에서 예상한 37%와 유사한 비율이며, 실험을 수행하기 좋은 결과이다.

실험은 Training Data를 기반으로 참조 모델을 생성하고, Test Data를 기반으로 예측 알고리즘에 적용한다. 참조 모델은 Bagging을 사용하여 Set1, Set2, Set3, Set4, Set5에 대한 Training Data Set과 Test Data Set을 추출하고 Training Data Set을 이용하여 생성한다. 예측 알고리즘은 Set1에 대한 Test Data를 적용하여 PV를 구한다. 그리고 PV(Prediction Variable)와 TV(Target Variable)를 비교하여 오차를 측정한다. 이와 같은 작업을 모든 Set에 반복 수행하여 실험 결과를 얻는다.

실험 결과 원본 데이터와 예측 데이터의 평균 오차율이 ±4.32%로 유효성이 높음을 검증하였다. 실험 결과를 살펴보면 대부분의 오차가 ±4.32% 사이에 있는 것을 확인할 수 있다. 그러나 비교적 큰 오차를 보이는 +18.69%과 -23.66% 같은 항목이 있다. 오차가 큰 항목의 원인은 설문 조사에서 프로젝트 특성의 분류 항목을 바르게 선택하지 못하거나, 프로젝트의 특성과 상관없이 확정된 예산에 맞추어 비용 산정을 시행한 경우이다.

4. 결론

프로젝트 비용 산정은 Top-down, 계수 모델, Bottom-up와 같은 기법으로 수행 하지만, 이후 협상에서의 비용 보정은 비기술적 요소들을 기반으로 수행되고 있다. 본 연구에서는 기술적 분석을 이용한 프로젝트 비용 보정 기법을 제안하고, 그 유효성은 실험을 통해 검증하였다. 검증 결과 실제 데이터와 예측 데이터의 평균 오차가 ±4.32%로 유효성이 높아서, 객관적이고 합리적인 비용 보정을 수행할 수 있다. 또한, Stakeholder 간의 비용 보정에 대한 견해차를 없애 신뢰성 향상과 상호 이익 극대화를 추구할 수 있다. 결국에는 프로젝트의 성공률 증대와 건전한 IT 산업의 발전 방향 등의 효과를 기대할 수 있다.

제안한 보정 기법은 Legacy 데이터의 기술적 분석을 판단하는 응답자의 능력에 따라 차이가 날 수 있으므로, 판단의 차이를 줄이기 위한 분류 항목들의 정량화에 대한 연구가 필요하다. 또한, 기술적 분석과 상관없는 비즈니스적 상황 때문에 예산이 미리 확정된 경우에 많은 차이가 날 수 있으므로, 예산이 확정된 프로젝트들을 분리하여 유효성을 높이는 방법에 대한 연구가 계속 되어야 할 것이다.

참고문헌

- [1] 한국소프트웨어진흥원 정제호, 소프트웨어 산업구조 선진화 방안, 정책연구 08-17, 2008.12.
- [2] The Standish Group International, CHAOS REPORT 2009.
- [3] Breiman, L. Bagging Predictors. Machine Learning, 24(2): 123-140, 1996.