# 過去の情報を用いたPBL向け工数見積り手法の提案

# 齋藤 尊 新美 礼彦 伊藤 恵

本研究では、学生主導で実施されるソフトウェア開発 PBL(Project-Based Learning)のマネジメント支援を目的とする。ソフトウェア開発経験の浅い学生がプロジェクトを管理する PBL では、学生の経験不足などが原因で工数見積りの誤りが発生し易い。一般的なソフトウェア開発では、開発工数を見積る手法を利用するが、データ収集とモデル構築・工数計算をプロジェクト進行中に行う必要があるため、学生が利用するのは難しい。一方 PBL ではプロジェクト管理ツールなどから様々な開発データを収集できるが、それらは十分に活用されていない。

そこで本研究では CoBRA 法をベースにした工数見積り手法を提案する.この手法では過去に実施された PBL の開発データを利用することによって,実施中のシステム開発 PBL の成果物作成に要する工数を見積る.本研究では見積り精度の評価と PBL 参加学生へのアンケートを実施するため,進行中の PBL への適用実験を行った.またこの適用実験では,データ収集の簡便化のためプロジェクト管理ツール運用ルールを設定した.

In this study, it is purpose that we support students to manage their PBL(Project-Based Learning). In PBL which students manage, delay of schedule and temporary tasks are easy to happen because of immature man-hour estimation. In a general software development, project members use man-hour estimation methods. But, it is difficult that participants use these estimation methods because they are busy to develop software. On the other hand, much data can be collected in PBL. However, it has not been used enough.

So, we propose the estimation method which is based on CoBRA(Cost estimation, Benchmarking and Risk Assessment). This method uses past PBL data and calculates man-hour it takes to implement their system. In this study, we conducted the experiment targeting PBL in progress in order to take a survey from participants.

# 1 序論

#### 1.1 背景

近年,実践的なソフトウェア開発教育手法の一つとして PBL(Project-Based Learning) が注目されている. PBL では学習者がチーム内の意思決定に基づき,問題発見・問題解決を行うことを重視した教育手法である [1]. 特にソフトウェア開発の分野では学習者が上流工程からソフトウェア開発を経験する場合が多

い. これは実際の社会が抱える問題に対する解決アプローチを経験できるという点で、特に注目されている. PBL によるソフトウェア開発は企業が事業として行う開発とは異なり、学習者がプロジェクトマネジメントや開発ツールの利用方法・開発言語を学びながらソフトウェアを開発する. そのため開発技術の不足から、プロジェクトのスケジュール遅延が発生しやすいと考えられる.

またソフトウェア開発において計画的なプロジェクト運行には、対象開発物への正確な開発工数の見積りが必要である。そのため、数学的な見積りモデル手法が複数提案されている。しかし、これらの手法は一定の知識や経験がなければ有効な利用が難しい。そのため、特に学生がプロジェクト管理を行うPBLでは利用が難しい。

Proposal of a Man-hour Estimation Method Using the Past PBL Data

Takeru Saito, Ayahiko Niimi, Kei Ito, 公立はこだて未来大学大学院システム情報科学研究科, Graduate School of Systems Information Science, Future University Hakodate.

コンピュータソフトウェア, Vol.29, No.1 (2012), pp.78-84. [研究論文] 2017 年 3 月 31 日受付.

# 1.2 研究目標

本研究では、PBLでは通常のソフトウェア開発と同様に複数の開発データを収集することができることに着目し、大学等の教育機関で保存している過去のPBL実績データを利用して進行中のPBLで行われるソフトウェア開発に要する工数を見積る手法を提案する。本手法はPBLに参加する学生が利用することを想定し、既存のプロジェクト管理ツールなどから収集できるデータのみを見積りに用いることで、PBL参加者に大きな負担をかけずに工数見積りを行うことを目指す。

また提案手法を評価するため,進行中の PBL を対象とした見積り実験を実施する.ここで算出した見積り結果を当該プロジェクトに開示し,見積り結果へのアンケート調査を行う.

### 2 関連研究

#### 2.1 PBL で発生する問題

永田らの大学院で実施された PBL では、スケジュール遅延などの問題が複数発生したことがわかる [2]. この報告では WBS(Work Breakdown Structure) によって実際の作業内容とその完了にかかる時間の見積りを行っていた. しかし、この見積りが正確に行えず、実際の工数との差異が大きくなったと報告している. この問題への対策として、直近の作業となるごとに詳細に予定工数を設定するローリングウェーブ計画法を提案している. しかし、この報告では直近の予定工数しか精密に見積りが行えないため、順次見積りを行わなければならず、PBL 参加者への作業量を増加させる問題がある. BPL は一般的に人員の増加がなされず、作業工数の増加がスケジュール遅延に直結する、そのため、PBL のスケジュール遅延にならないよう、極力作業工数を増やさない見積り手法が必要である.

#### 2.2 見積り手法

ソフトウェア開発における工数見積りは一般的に過去の類似プロジェクトの実績を参考に見積る類推法や専門家による判断,プロジェクト特性を説明変数として用いて工数を見積るパラメトリック法がある

# 2.2.1 ファンクションポイント法

ソフトウェア開発における工数見積りを行う場合,開発対象の開発規模を推定することが必要である場合が多い.開発規模の代表的な指標としてファンクションポイント (FP) がある[4]. FP は上記区分ではパラメトリック法に属する方法で導出されるが,この手法は開発工数そのものではなく,開発対象の規模を算出する.

#### 2.2.2 説明変数を用いる工数見積り手法

ソフトウェア開発プロジェクトには、開発規模だけでなく開発環境や開発にあたって遵守しなければならない条件、参加メンバーの習熟度など複数の要因が工数に影響を与える。そのため説明変数を用いて工数を算出する場合、FPだけでなくそれらの要因を説明変数として利用しなければならない。この説明変数には開発規模やプロダクト、人材に関するメトリクスなどが一般的に必要とされている[5]。これらの説明変数のうちいずれを利用するかは見積り手法やプロジェクトによって変わる。

#### 2.3 PBL を対象とした工数見積り

PBL を対象とした見積り手法に関して八島らはユースケースポイント (UCP) 法による見積りの検証を行っている[6]. この研究では計画段階で作成されたユースケースでは見積り精度が悪いという問題点がある. UCP 法は FP 法と同様, 開発規模や作業量だけの算出であるため, 工数見積りに直接利用するには不十分である.

#### 3 提案する見積り手法

本研究では過去に実施された PBL の実績データを利用し、見積りモデルの構築と見積り結果の算出を行う。そのため、CoBRA 法を元にした工数見積り手法を提案する。本章ではまず CoBRA 法について述べ、その後、提案手法について述べる。

### 3.1 CoBRA法

CoBRA 法 (A Hybrid Method for Cost Estimation, Benchmarking, and Risk Assessment) とは, 説明変数を用いる見積り手法のうち, 過去に実施され

たソフトウェア開発の定量的データとソフトウェア開発熟練者の経験をもとに見積りを行う手法である[7]. CoBRA 方では以下 3 つを前提条件としている.

- 1. ソフトウェア開発熟練者は開発コストを,経験をもとに定量的に把握できる.
- 2. 純粋状態において工数と開発規模は比例する.
- 3. 開発における諸要素によって,工数と開発規模 の線形関係に振れが生じる.

前提条件 3 より,開発期間に影響を与える要素を工数変動要因と呼ぶ.この工数変動要因は  $CO_i$  と表す.また工数変動要因による工数への影響が全く無く,前提条件 2 のみが見積り結果に現れる状態を本研究では純粋状態と呼ぶ.

この CoBRA 法では過去のプロジェクト実績データを利用し、見積りモデルを構築するため本研究の目的に適している.しかし、PBL では CoBRA 法のモデル構築に必要な見積り熟練者の協力を得ることが難しい.そのため、本研究では見積り熟練者の協力がなくても見積りモデルが構築できるよう、CoBRA 法を改変した手法を提案する.

### 3.2 提案手法

3.1 節より、見積り熟練者の協力を得ずに見積りが可能となるよう、通常の CoBRA 法の前提条件 3 つに加えて以下 2 つを前提条件に追加する.

- 4. 見積り対象プロジェクトが開発期間中に予想した開発工数は純粋状態の見積り結果と一致する.
- 5. *CO*<sub>i</sub> は当該工数変動要因の取得値と,実績工数 を利用して算出した重回帰係数の積に比例する.

前提条件 4 は PBL 参加者は工数見積りの際,工数変動要因を考慮できていないことが多いため設定した。この前提条件 5 により,前提条件 1 の開発熟練者の協力がなくとも  $CO_i$  の値を収集データからのみ算出できる.

これらの条件を踏まえ、本研究では以下の見積りモデルを提案する.

$$PM = \alpha \times Size \times (1 + \beta \times \sum (Data_i \times Coeff_i))$$
(1)

式1のPMは算出する見積り工数を表し、Sizeは

開発規模を表している。また各工数変動要因の取得値が  $Data_i$ , その重回帰係数が  $Coeff_i$  である。  $Data_i$  に該当する数値として,プロジェクト参加人数などがある。加えて  $\alpha$  は開発規模 1 単位あたりの必要工数, $\beta$  が  $Data_i \times Coeff_i$  を工数変動要因の影響率とするための調整定数である。

この見積りモデルによって見積りを行う際には、まずモデル構築によってパラメータ $\alpha$ 、 $Coeff_i$ を設定する。その後、見積り対象のプロジェクトのデータを収集し、当該プロジェクトの開発規模及び各工数変動要因の取得値をSize、 $Data_i$  に代入し計算する。この見積り手法では当該プロジェクトが開発するシステムの実装期間を見積り結果として出力する。ここで算出する見積り結果は実際に開発を行っている時間ではなく、開発を行っていない時間も含めた工数を意味している。これはPBL に参加する学生は演習時間外にもPBL 活動を行っており、学生の作業負担を増やさずに正確な作業時間を計測することは困難であることが理由である。

#### 3.2.1 モデル構築手順

式1は以下の手順で構築する.

1. データ整形

収集した各データのうち、欠損値はその欠損部分に 0を代入する。また全プロジェクトで値が欠損している工数変動要因は、その要因自体をデータセットから除外する。

すべての工数変動要因を数値として利用するため,名義尺度に属する工数変動要因をダミー変数に変換する. また多重共線性を排除するため,各工数変動要因間の分散拡大係数 (VIF) が 10 以上となる組み合わせのうち 1 種を除きすべて除外する.

2. αの算出

3.2 節の前提条件 4 より,目的変数を Expected,説明変数を Size とした最小二乗法により定数  $\alpha$  を算出する.

3. Coeff<sub>i</sub> の算出

手順1で整形した工数変動要因を元に、 $Coeff_i$ の 算出を行う。実績工数を Actual としたとき,前 提条件 3 及び 4 より  $Coeff_i$  は実績工数と当該プ ロジェクトの予想工数の差と等しい. そのため以 下の式 2 に対し目的変数を Actual - Expected, 説明変数を Data<sub>i</sub> とした重回帰分析を行う.

 $Actual - Expected = \sum Data_i \times Coeff_i$  (2) この重回帰分析をもとに算出した標準化偏回帰 係数を  $Coeff_i$  に設定する.

#### 4. βの算出

上記の手順によって算出した lpha, $\mathit{Coeff}_i$  を説明 変数, Actual を目的変数とした回帰分析を行う. そこで得られた回帰係数を $\beta$ に設定する.

# 3.3 工数変動要因とその他見積り用パラメータの 収集

#### 3.3.1 対象プロジェクト

本研究では見積りモデル構築や見積り実験実施のた め, 著者ら所属大学で 2013 年から 2015 年に実施さ れた PBL29 件からデータを収集した. そのうち 11 件はおおよそ1週間程度の期間で実施される短期集 中の PBL で、通常の PBL と異なり 1 日の講義時間 枠のすべてを実施時間に利用する PBL であるため, 4章の実験では残り18件のデータを見積りモデル構 築に利用した.

#### 3.3.2 メトリクス

対象プロジェクトから工数変動要因を42種,その他 見積り用パラメータを4種収集した. モデル構築用パ ラメータとして予定工数や FP, 工数変動要因としてプ ロジェクト参加者数やユースケース数、開発言語など を収集した. これらは議事録やプロジェクト管理ツー ルの記録、ソースコード等から算出した. 収集したメト リクスの種類は https://github.com/keilab2016/ PBLEstimationTool/blob/master/palam-def.md, そ 本研究で利用するデータを確実かつ容易に収集でき の収集結果は https://github.com/keilab2016/  ${\tt PBLEstimationTool/blob/master/Teaching\_data}.$ csv にて確認できる. このメトリクスのうち、議事録 等で記録されていないなどの原因でデータが欠損し ていた場合は NA を入力した.

#### 4 進行中 PBL への適用実験

本章では 2016 年度に実施された PBL を対象とし た見積り実験について述べる.



図 1 2016 年度 enPiT 分散 PBL

#### 4.1 実験概要

# 4.1.1 見積り対象 PBL

この実験を行うため、2016 年度に enPiT で行われ た PBL のデータの収集を行った. enPiT とは "大学 間・企業間で連携を取りつつ ICT を用いた新たな価 値創出ができる人材の育成"を行っている事業である [8]. 本研究では enPiT において著者ら所属大学で行 われた分散 PBL(図 1)6 チームのデータを見積り実験 の対象とした.

#### 4.1.2 実験準備

本研究はこの PBL の開始時点から、データ収集 に向けた準備を行った. この PBL ではプロジェク ト管理および構成管理のために GitHub を利用して いる[9]. そこでこの GitHub からデータ収集を行っ た. また本実験を実施するに当たって, 必須データ の確実な記録とデータ収集の簡略化のため、議事録 フォーマットの利用や Issue 発行に関するルールなどの GitHub 運用ルールを設定した. 設定した GitHub 運用 ルールの詳細は https://github.com/keilab2016/ PBLEstimationTool/blob/master/Operational\_rule. md に記載する. これらの GitHub 運用ルールを設定 することで、ルール設定を受けていない PBL に比べ、 るようになった.

#### 4.1.3 収集データ

収集できたデータは https://github.com/ keilab2016/PBLEstimationTool/blob/master/2016\_ enPiT.csv にて確認できる.表1は収集データの抜 粋である.

表 1 収集データ例

	Estimated	Actual	FPTrial	FPApproxi
2016_a	144	116	330	103
2016_b	216	126	85	44
2016_c	144	172	135	67

#### 4.2 実験結果と考察

4.1.3節の PBL データを用い、見積り実験を行った。また見積りにおいて、3.3.1節のデータを見積りモデル構築に用いた。モデル構築のため手順1のデータ整形を行った結果、18種の工数変動要因が見積りに適用された

この見積り結果が表2である.ここで表2の見積り工数及び予想必要日数は全プロジェクトのデータ収集が完了した直後に算出し、実績工数及び相対誤差は全チームが開発対象システムの実装を終えて以降、データ収集と算出を行った.表2のうち、プロジェクト2016-aの見積り工数は436.4151人日となった.表2の相対誤差平均は116.2700%、中央値は9.8042%となった.この結果より、外れ値による見積り精度への影響が大きいことが確認できた.特にプロジェクト2016-aの相対誤差が非常に大きいが、これは参照ファイルの量によって開発規模に利用したFP値が実態よりも過剰に大きくなったことが原因と考えられる.

2016 年 11 月 25 日に行われた PBL 講義時間に、この表 2 の見積り工数及び予想必要日数を当該プロジェクトの参加者に伝えた. この時点でプロジェクト C, E, F はすでに開発開始から 3 週間以上が経過しており、その他は開発開始から 2 週間以内であった.

# 4.3 アンケート調査

本見積り結果に関して PBL 参加者に対し、見積り結果と結果の開示時期及び GitHub 運用ルールに関してアンケートを実施した。このアンケートでは、発表した見積り結果を有用に活用できたのか、運用ルールにかかるコストが PBL で許容できるものであったかなどを調べるため実施した。

### 4.3.1 アンケート項目と結果

アンケートを 2016 年 12 月 9 日に実施し, その結果 29 名中 19 名から回答を得られた.

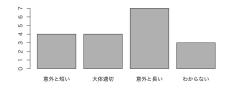


図2 設問1の回答

設問1"お伝えした見積り結果はあなたの想像と比 べてどうでしたか"の回答結果は図2の通りである. この設問は見積り結果がどの程度の精度であると感 じたかを調べるため実施したが、図2より"意外と 長い"が最も多い回答だった. 設問 2"見積り結果は マネジメントに関して有用につかえましたか"という 質問に対し、"はい"という回答が3件、"いいえ"が 16件となった.この設問は見積り結果がプロジェク トで実際に活用されたか否かを調べるため実施した. 設問3では設問2の回答結果に関して理由を質問し た. その結果, "気にしていなかった" という意見が 最も多かった、また設問 4"見積り結果の発表タイミ ングはどうでしたか"の回答として, "わからない" が9件, "ちょうどいい" が5件, "遅すぎる" が5件 となった. これらに加え、GituHub 運用ルールに対 する感想をアンケートにて収集した.

#### 4.3.2 アンケート結果の考察

4.3.1節の結果より,各回答をチームごとに分割して検証した.その結果,各チームごとに設問 4への回答に傾向が見られた.例えばチーム D とチーム E のメンバーのうち 3 分の 2 が "遅すぎる" と回答しており,逆にチーム E はメンバー全員が "ちょうどいい" と回答していた.これは見積り結果発表時点での各チームの進捗状況の差が原因と考えられ,実際にチーム  $D \cdot E$  で "遅すぎる" と回答したメンバーの過半数は,設問 5 で "見積り結果がほしい時期とずれていた"という旨の記述を行っていた.この実験では全チームのデータを集め終わった段階で見積りを行った.しかし,今後見積りを行う場合は,データが収集できたチームから順次見積りとその結果の通知を行う必要があることがわかった.

見積り結果が有効に使えたかの理由を聞いた設問3 では、本来どのような目的で見積り結果を利用したか

表 2 適用実験の見積り結果

プロジェクト	2016_a	2016_b	2016_c	2016_d	2016_e	2016_f
見積り工数 (人日)	436.4151	160.1002	185.6781	96.0558	234.4776	80.7914
実績工数 (人日)	116	126	172	140	210	175
相対誤差 (%)	276.2199	27.0637	7.9524	-31.3887	11.6560	-53.8335

を確認するため実施したが、回答結果では"気にしていなかった"が最も多くなった。同様に結果開示のタイミングが早かったか遅かったかを聞いた設問 4 では、"わからない"が最も多くなった。これらの結果より今回の実験では多くの参加者が見積り結果に関してあまり関心を示していなかったことが分かった。そのため、特に PBL を対象に見積りを行う場合、見積り結果の扱い方に関して事前により詳しく説明をする必要がある。

また GitHub 運用ルールのうち "Issue 発行ルール" が最も手間になったと回答され、この回答の半数以上が "他の方法で情報共有ができていたから不要だった" という旨を回答していた.そのため、データの収集方法や PBL の運用方法を改めるなどの対策を行う必要があると考えられる.

#### 5 結論

#### 5.1 まとめ

本研究は CoBRA 法をベースにした PBL 向け工数見積り手法を提案し、進行中 PBL への適用実験を行った。実験では必要データの収集を用意にするためプロジェクト管理ツールの運用ルールを指示してPBLを進行させた。結果対象プロジェクトの多くでは一定の精度で工数を見積ることができた。しかし、プロジェクトごとに開発の進捗に違いがあったため、適切なタイミングで見積り結果を提供できないプロジェクトが複数出てしまった。また算出した見積り結果の扱い方について説明が足らず、有効に活かされなかった。だが工数変動要因を収集するために設定したGitHub運用ルールによって、情報が見つけやすくなったなどの意見が得られた。

#### 5.2 今後の展望

本研究では、PBLのマネジメント支援のためにどの程度の見積り精度が必要であるかを調べられていないため、実施中のPBLを対象としたさらなる評価実験が必要である。またPBLが担当教員や参加者が利用できるように提案手法のツール化が必要である。

#### 参考文献

- [1] "PBL を基軸とする工学教育プログラム | 九州工業大学." [Online]. Available: http://www.mns.kyutech.ac.jp/~nakao-m/pbl/about.html. [Accessed: 28-Dec-2016].
- [2] 永田佑輔 and 山戸昭三, "1706 PBL で発生した 問題とその解決事例 (一般セッション)," プロジェクト マネジメント学会研究発表大会予稿集, vol. 2011, pp. 504-506, Mar. 2011.
- [3] 松澤芳昭, 塩見彰睦, 秡川友宏, and 酒井三四郎, "ソフトウェア開発の教員主導型 PBL における反復プロセスと EVM 導入の効果," 研究報告コンピュータと教育(CE), vol. 2009-CE-99, no. 9, pp. 1-8, May 2009.
- [4] 岡村正司, "ファンクション・ポイント (IFP)," in プロジェクトコスト見積もり入門: ファンクション・ポイント、COCOMOII、WBS によるソフトウエア開発コストの導き方, 1st ed., 日経 BP 社, 2009, pp. 87-116.
- [5] "エンタプライズ系事業/見積もり手法: IPA 独立行政法人 情報処理推進機構." [Online]. Available: http://www.ipa.go.jp/sec/ softwareengineering/std/ent01-c.html. [Accessed: 25-Dec-2016].
- [6] 八島亮平, 山戸昭三, and 和田耕一, "1202 Project Based Learning における見積もり手法に関する考察: ユースケースポイント法の適用 (一般セッション)," プロジェクトマネジメント学会研究発表大会予稿集, vol. 2011, pp. 99-102, Sep. 2011.
- [7] Adam Trendowicz, Jens Heidrich, Jurgen Munch, Yasushi Ishigai, Kenji Yokoyama, Nahomi Kikuchi, "Development of a hybrid cost estimation model in an iterative manner," ICSE 2006: 331-340.
- [8] "enPiT 分野・地域を越えた実践的情報教育協働ネットワーク." [Online]. Available: http://www.enpit.jp/. [Accessed: 15-Dec-2016].
- [9] "Build software better, together," GitHub. [Online]. Available: https://github.com. [Accessed: 29-Dec-2016].
- [10] 緒方裕光 and 柳井晴夫, 統計学:基礎と応用, 初版. 京都: 現代数学社, 1999.

# 齋 藤 尊

2015 年公立はこだて未来大学システム情報科学部情報アーキテクチャ学 科卒. 2017 年公立はこだて未来大学 大学院システム情報科学研究科博士

前期課程修了. 同年富士通株式会社入社.

# 新美 礼彦

2002 年桐蔭横浜大学大学院博士後期 課程制御システム工学専攻修了(博 士(工学)). 同年公立はこだて未来 大学システム情報科学部助手. 2009 年より同大学准教授. データマイニングに関する研究 に従事. IEEE-CS, IEEE-SMC, 情報処理学会, 人工 知能学会, 日本知能情報ファジィ学会各会員.

# 伊藤恵

1998 年北陸先端科学技術大学院大学 情報科学研究科博士後期課程修了. 同 年同研究科助手. 2001 年公立はこだ て未来大学講師. 2013 年同大学准教

授.ソフトウェア科学会,情報処理学会,教育システム情報学会,観光情報学会,IEEE CS 各会員.