

Base Functional Components を用いた工数見積り精度向上について

天 寄 聡 介^{†1}

ソフトウェアの規模は工数見積りにおいて最も重要な予測因子である。規模は FP や LOC などの単一のメトリクスで記録することが多い。近年、FP の構成要素である BFC を予測因子とすることで工数見積り精度の向上を試みる研究が報告されている。しかし、これらの研究で用いられている工数見積り手法は線形回帰モデルに限られており、他の手法における有用性が確認されていない。本研究では類推法における BFC の有用性について実証実験をおこなった。本ワークショップでは、BFC を工数見積りに用いることの有用性について議論したい。

Improving Effort Estimation Accuracy with Base Functional Components

SOSUKE AMASAKI^{†1}

Product size is the most important factor for effort estimation. FP and LOC are major metrics for sizing. Recent studies reported that using Base Functional Components instead of FP may improve predictive accuracy. However, these studies did not consider effort estimation methods other than linear regression models. In this paper, we show the result of using Analogy-based method (ANN) with BFC metrics. In the workshop, we would like to discuss regarding the usefulness of BFCs-based effort estimation.

1. はじめに

工数見積りモデルに関する研究が盛んに行われている¹⁾。ソフトウェアの規模は工数見積りにおいて最も重要な予測因子である。FP や LOC などの単一のメトリクスで規模を記録することが多い。近年、FP の構成要素である BFC を予測因子とすることで工数見積り精度の向上を試みる研究が報告されている²⁾。しかし、これらの研究で用いられている工数見積り手法は線形回帰モデルに限られており、他の手法における有用性が確認されていない。本研究では類推法における BFC の有用性について実証実験をおこなった。

2. 実 験

工数見積りの研究で使用されるデータセットには PROMISE Repository で公開されているものや ISBSG Group が提供しているものなどがある。これらデータセットの多くでは規模が FP で計測されている一方で、BFC まで記録しているデータセットは少ない。そこで、本研究では ISBSG R11 database³⁾ を用いる。ISBSG データセットから実験に使用するデータセットを表 1 の手順に従って抽出した。

表 1 実験データセットの抽出手順

	属性	基準	Remained
0	—	—	5052
1	Data Quality	= {A B}	4744
2	Summary Work	= Norm. Work	3142
3	Resource Level	= 1	2679
4	FP 手法	≥ IFPUG 4	723
5	\sum BFCs	= FP	256

類推法ではユークリッド距離で類似性をはかり、見積もりには近傍の工数の平均値を用いた。各説明変数は見積もりに際して正規化される。近傍数は $k = 1, 2, 3, 5$ で実験を行った。また、BFC の選択には wrapper approach⁴⁾ を用いた。最適化する評価基準には MdmRE を使用した。

工数見積りモデルでは工数および規模メトリクスの対数を用いることが多い。分布の対称性が増し偏差が減少するため、対数化によって見積り精度の改善が見込まれる場合がある。しかし、BFC を用いた場合はどの要素を対数化することがよいのか明らかでない。一方、類推法では説明変数間の相関を考慮しなくてもよい利点がある。これらを踏まえて、以下の 4 パターンで工数見積り精度を比較した。

^{†1} 岡山県立大学

Okayama Prefectural University

表 2 実験結果 (MMRE)

変数の扱い	$k = 1$	$k = 2$	$k = 3$	$k = 5$
Effort, FP	1.36	1.38	1.39	1.37
log(Effort), log(FP)	1.32	1.32	1.32	1.32
(1)	1.77	1.39	1.36	1.31
(2)	1.56	1.24	1.17	1.12
(3)	1.55	1.50	1.46	1.43
(4)	1.52	1.27	1.17	1.05

- (1) 工数: 前処理なし, BFC: 前処理なし
(2) 工数: 対数化, BFC: 対数化
(3) 工数: 前処理なし, BFC: 対数化, 前処理なし
(4) 工数: 対数化, BFC: 対数化, 前処理なし

実験は 10 x 10-fold cross-validation⁵⁾ で行い, 各パターンについて最も見積もり精度が高い BFC の組み合わせを選ぶ。10 x 10-fold CV の結果全体で評価するため, 実際の予測でも同じ BFC の組み合わせがよい精度を示すと考えられる。この結果を FP を用いた場合の見積もり精度と比較することで BFC の有用性を評価する。

3. 結果

実験結果を表 2, 3 に示す。FP を説明変数とした場合, 近傍数の変化によらず MMRE, MdMRE に大きな変化はない。また, 対数化によってわずかながら MMRE が向上しているが, MdMRE は低下している。Wincoxon の符号付き順位検定を用いると, $\alpha = 0.05$ で統計的に有意差があるのは MMRE の場合で $k = 1$ のみであった。MdMRE の場合, $k = 3, 5$ のとき統計的な有意差が認められた。以上から, $k = 3$ で対数化しない場合が最も見積もり精度が高いといえる。

BFC を説明変数とした場合, 近傍数によって MMRE, MdMRE とともに大きな変化がみられる。MMRE, MdMRE とともに, 工数を対数化した (2), (4) が (1), (3) より精度が高い。(1) と (2) の間では $k = 2, 3, 5$ で統計的有意差があった。また, (3) と (4) の間では $k = 2, 3, 5$ でほとんどの場合で有意差があった。(2) と (4) の間では, $k = 5$ の MMRE でのみ有意差があった。以上から, $k = 5$ の (4) が最も見積もり精度が高いといえる。

$k = 1, 2$ の時の MMRE 以外では, (4) は FP を用いた場合より統計的に見積もり精度が高かった。このことから, 類推法において BFC を用いることが精度の上で有用である可能性が確認できた。

4. 考察

BFC を用いた (4) で $k = 5$ のときに選択されたメ

表 3 実験結果 (MdMRE)

Model	$k = 1$	$k = 2$	$k = 3$	$k = 5$
Effort, FP	0.70	0.68	0.65	0.66
log(Effort), log(FP)	0.70	0.70	0.70	0.70
(1)	0.67	0.62	0.59	0.59
(2)	0.65	0.56	0.55	0.53
(3)	0.65	0.59	0.57	0.56
(4)	0.65	0.57	0.55	0.52

トリクスは次の通りである。

- 対数化したもの: EI, EO, EQ, ILF
- 対数化していないもの: EI, ILF

EIF が含まれない一方で, EI, ILF は対数化したもの, していないもの両方が含まれている。また, (2) で $k = 5$ の場合に選択されたメトリクスは, 上記の対数化したものと一致した。つまり, MMRE における有意差は EI, ILF の有無が影響しているといえる。また, EI, ILF の類似性が他の BFC より重要である可能性を示している。以上のことから, BFC の種類によって工数に及ぼす影響が異なっており, そのことが見積もり精度の向上につながったと推測できる。

5. おわりに

本稿では BFC を規模メトリクスに用いることが類推法による工数見積もり精度の向上に有効である可能性を示した。本実験の結果は最良のモデルの場合であり, 実際の場面で適切な BFC を選択できるかについては明らかでない。この点は今後の課題である。

参考文献

- 1) Jørgensen, M. and Shepperd, M.: A Systematic Review of Software Development Cost Estimation Studies, *IEEE Trans. Softw. Eng.*, Vol.33, No.1, pp.33–53 (2007).
- 2) Gencel, C. and Buglione, L.: Do Base Functional Component Types Affect the Relationship between Software Functional Size and Effort?, *Proc. of Mensura 2007*, pp.72–85 (2007).
- 3) International Software Benchmarking Standards Group (ISBSG): ISBSG Estimating, Benchmarking and Research Suite Release 11 (2004).
- 4) Dejaeger, K., Verbeke, W., Martens, D. and Baesens, B.: Data mining techniques for software effort estimation: a comparative study, *IEEE Trans. Software Eng.* (2011).
- 5) Bouckaert, R.R.: Choosing between two learning algorithms based on calibrated tests, *Proc. of ICML 2003*, pp.51–58 (2003).