

工程別欠陥埋め込み数の予測

三菱電機マイコン機器ソフトウェア（株）

衣簀宏和

ki-hiro@mms.co.jp

品質管理における問題点

ソフトウェア品質管理において、工程別の品質を把握したいというニーズがある。その際、レビューでの検出件数に加えて、当該工程で欠陥がどの程度埋め込まれたかを併せて検討する事が精度の良い品質管理の為に必要である。

手法・ツールの適用による解決

SWAD※, SWDD※, CD※の各工程を対象に定量的要因に加えて、定性的要因を考慮し、実データに基づいた欠陥埋め込み数の予測モデルの構築を行い、構築したモデルを評価した。

手法とツール:

Step-Wise変数選択, 重回帰分析, leave-one-out交差検証, MRE※

※SWAD=ソフトウェア方式設計, SWDD=ソフトウェア詳細設計, CD=コード作成, MRE=Magnitude of Relative Error

分析対象データと分析手順

分析データには、2008年-2011年までの実プロジェクトデータを用いた。各プロジェクトでは開発タスク毎に工程単位で欠陥埋め込み数, 定量的データ, 定性的データを収集している。

以下の手順で分析と評価を実施。

Step1: 各メトリクスの分布と傾向の確認と対数変換

※開発規模と欠陥埋め込み数は偏りのあるデータであった。データ正規分布として評価したい為、自然対数変換を行った。(LNを付与して表記)

Step2: 外れ値・欠損データの除外

Step3: 散布図, 相関係数, 経験的知識から説明変数を選定

※表1に選定した説明変数の候補を示す。

Step4: Step-Wise変数選択による重回帰分析

Step5: leave-one-out交差検証とMRE分布によるモデル評価

表1. 説明変数の候補とLN欠陥埋め込み数との相関係数

メトリクス	記号	SWAD	SWDD	CD
LN_開発規模	—	0.76	0.92	0.81
開発対象に関する事前勉強の有無	M1	-0.32	—	—
要求性能の難易度	M2	-0.49	-0.55	-0.30
システム構造の複雑度	M3	-0.53	-0.42	—
既存システムからの流用の可否	M4	-0.48	-0.52	-0.36
仕様変更の可能性	M5	-0.44	—	—
未解決課題の数	M6	-0.29	—	—
直近の使用変更の頻度	M7	-0.32	—	—

Step-Wise変数選択による重回帰分析

工程	説明変数	偏回帰係数(P値)
SWAD	定数項	3.86(0.000)
	LN_開発規模	0.52(0.000)
	M5	-0.43(0.004)
	調整済みR2	0.675
	N	26
SWDD	定数項	2.23(0.000)
	LN_開発規模	0.99(0.000)
	M2	-0.29(0.153)
	調整済みR2	0.847
	N	26
CD	定数項	1.71(0.000)
	LN_開発規模	0.820(0.000)
	調整済みR2	0.64
	N	24

MREによるモデル評価

選択された説明変数を用いてleave-one-out交差検証を行った。予測誤差の評価ではMRE((予測値-実績値)÷実績値)を用いた。

SWADの予測精度は、実用として十分な精度があるが、SWDD及びCDはSWADと比較して分布のブレ幅が大きく精度としては十分とは言い難い。

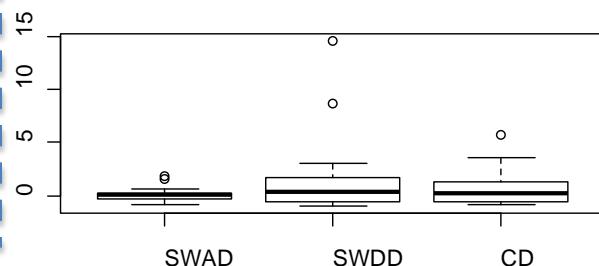


図1. 工程別のMRE分布

結果と課題

■結果

- 実データから予測モデルを構築・評価するノウハウを獲得。
- 予測モデルから所属部門の欠陥の埋め込み数削減の注力すべきポイントが見えてきた。

■課題

- SWDD及びCDのバラつき発生の原因を調査する必要がある。
- 外れ値の発生原因の調査により各予測モデルを見直す必要がある。
- 得られたモデル式の実務への適用方法・見直しを考えていく必要がある。