



Empowered by Innovation **NEC**

プロセスメトリクスを用いた システムテスト検出バグ数の予測

2012/9/13

日本電気株式会社 高橋 亮介

目次
1. 現在の課題と本研究の目標
2. 開発の特徴と開発プロセスについて
3. 分析対象データの概要
4. 目的変数の設定
5. 相関分析による評価
6. 重回帰分析によるモデル式の作成
7. 作成したモデル式の評価
8. 考察
9. 今回の成果と今後の課題

Page 2 © NEC Corporation 2012 Empowered by Innovation **NEC**

1. 現在の課題と本研究の目標

現在の業務と課題

現在の業務

- ソフトウェア製品品質保証(SQA)
第三者の観点でメトリクスを用いたソフトウェアの開発/品質管理や
成果物のレビュー、第三者評価を実施しソフトウェアの出荷品質を保証する

開発プロジェクトとしての課題

- 出荷納期を遵守できない開発プロジェクトがある
 - 開発終盤のテスト(弊社ではシステムテスト)で多くのバグが発生し、
バグ対応に多くの工数を費やしてしまう
(開発終盤のため後戻り工数増大)

ソフトウェア製品品質保証部門としての課題

- システムテストで検出されるバグ数に焦点を絞った見積りモデルがない
- 現在測定しているメトリクスから開発状況の異常(システムテストで
バグが多発する要因)を検知できない

本研究の目標とアプローチ

目標

- システムテスト(以下ST)のバグ数を見積もるための**モデル式を作成**
- 作成したモデルに基づいて、**STバグ数を削減するための有効な手段**(STバグ数と強い因果関係にあるメトリクス)を見出す

アプローチ

- 現在測定しているメトリクスに対して多変量解析を実施
 - STで検出されたバグ数とその他メトリクスとの関連性(相関関係)を分析した上で、回帰分析を実施しモデル式を作成する
- 多変量解析に用いるツールはJUSE-StatWorksを使用

2. 開発の特徴と開発プロセスについて

ソフトウェア製品の担当領域と開発の特徴

HW部門とのシナジーによる強いプラットフォーム製品事業を推進



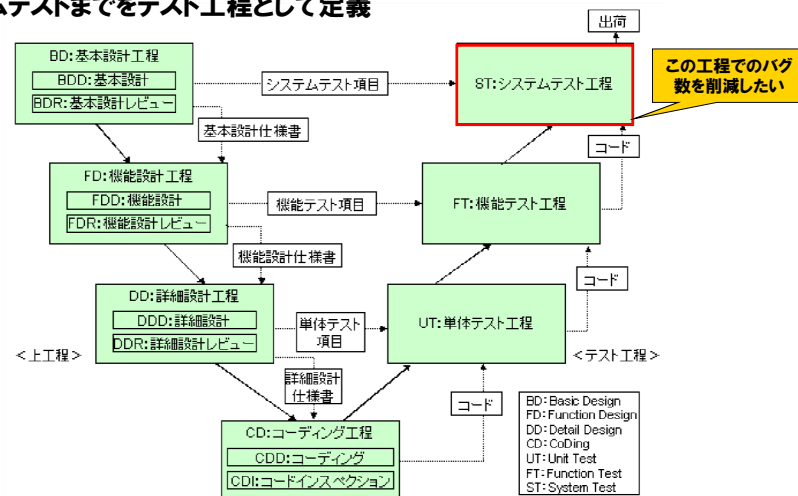
担当事業部の開発の特徴

- 特定の顧客向けではなく、汎用的なソフトウェアの開発が多い
- 新規製品の開発は少なく、既存製品に対する機能強化(バージョンアップ)を行う開発が多い

開発プロセス

基本的にV字のウォーターフォールモデルを採用

基本設計からコーディングまでを上工程、単体テストからシステムテストまでをテスト工程として定義



3. 分析対象データの概要

分析データの概要

■ 開発期間:2010年度～2011年度に実施した開発

■ データサンプル数:302個

■ 開発規模が10LOC～数万LOCと非常に小規模な開発から大規模な開発のデータ
※弊社では自分達で作り込んだ開発規模を自製規模と呼ぶ

メトリクス項目一覧

(1) STバグ数	(11)各工程の設計製造工数(BD～CD)
(2) 開発規模 ※自製規模+流用規模	(12)各工程のレビュー工数(BD～CD)
(3) 自製規模	(13)上工程レビュー比率
(4) 上工程バグ数	(14)上工程レビュー比率(CD工程除く)
(5) 全体工数(ST除く)	(15)各工程のレビュー比率(BD～CD)
(6) テスト工程(ST除く)	(16)作り込み品質 ※上バグ/上レビュー工数
(7) 上工程設計製造工数	(17)全体仕様書ページ数
(8) 上工程レビュー工数	(18)各工程の仕様書ページ数(BD～CD)
(9) 上工程設計製造工数(CD除く)	(19)各工程の新規テスト項目数(UT,FT,UT+FT)
(10)上工程レビュー工数(CD除く)	

■ 上記メトリクス項目一覧の内、(1)～(4)と(17)～(19)を**プロダクトメトリクス**、
残りの項目を**プロセスメトリクス**と呼ぶこととする

■ (1)は目的変数、(2)～(19)は説明変数の候補となる項目とする

4. 目的変数の設定

目的変数の検討

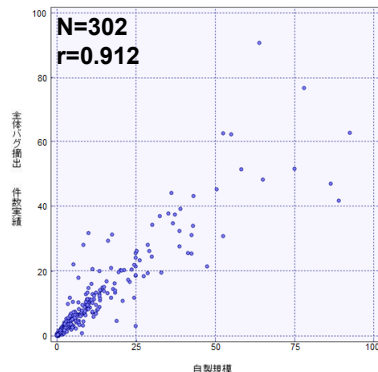
- 見積もりモデルを作成する上で目的変数となるのはSTバグ数である
- 一般的には自製規模とバグ数は相関関係が強いと考えられる
- 以下の2つのどちらを目的変数として設定すべきか検討する

1. STバグ数を自製規模で割った**STバグ密度**
2. 自製規模で割らない**STバグ数**

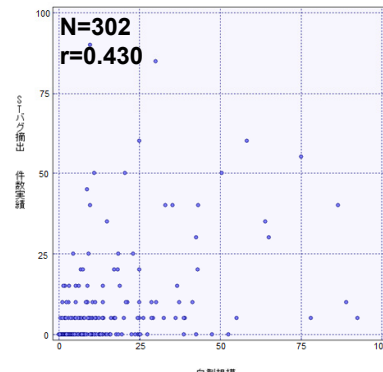
■ アプローチ

- 自製規模が大きければ大きいほどバグ数も多くなる傾向が当てはまるのは、全体バグ数とSTバグ数の両方に言えるのかを分析する
- 以下の2つの相関係数を比較してみる
 1. 自製規模と全体バグ数
 2. 自製規模とSTバグ数

自製規模とバグ数の散布図と相関係数



自製規模と全体バグ数の散布図



自製規模とSTバグ数の散布図

注)図で表示している目盛りの値は、ある数値で割った値を100としている

自製規模と全体バグ数の相関係数は0.912

自製規模とSTバグ数の相関係数は0.430

目的変数の決定

結論

- STバグ数は全体バグ数と異なり自製規模の影響を強く受けない
- STバグ密度を目的変数として分析すると、分析結果でデータにばらつきがあった場合、STバグ数の影響でばらついているのか、自製規模の影響でばらついているのかの判別が難しくなる

目的変数の決定

- 分析に用いるデータはSTバグ数を自製規模で割ったSTバグ密度を使用せず、**STバグ数**を目的変数として使用する

5. 相関分析による評価

STバグ数と各項目との相関分析

■ モデル式を作成する上で、まずはSTバグ数と各項目との相関関係を分析して、STバグ数と影響が強そうな項目を探す

■ 相関分析の結果、STバグ数と各項目との相関係数は以下の表の通りだった
(特に相関が強かった5つの項目のみ表示)

項目名	相関係数
全体工数実績(ST除く)	0.488
上工程設計製造工数	0.501
FD設計製造工数	0.492
DD設計製造工数	0.467
CD設計製造工数	0.520

■ 一番強い相関係数はCD設計製造工数の**0.520**であり、一般的に強い相関と言える0.7以上の相関係数がある項目は見当たらなかった

分析対象データの絞込み

■ 全データを分析対象とするとSTバグと強い相関がある項目は見当たらないため、データの絞込みを検討する

■ 一般的に大規模な開発ほど開発コストがかかり、重点的に開発管理を行なう対象となり得ることから、小規模な開発のデータは除外して、ある程度
自製規模の大きいデータに絞って分析することにする

■ 今回は自製規模を以下の3つに分けてデータ分析を実施する

1. 自製規模1K LOCより大きい開発
(データ数199個、全データ数に対する比率65.9%)
2. 自製規模5K LOCより大きい開発
(データ数78個、全データ数に対する比率25.8%)
3. 自製規模10K LOCより大きい開発
(データ数34個、全データ数に対する比率11.3%)

絞り込んだデータによる相関分析

■ 1K LOC、5K LOC、10K LOCの3つに分けたデータを分析してみた結果、STバグ数と一番強い相関係数はそれぞれ以下の値となった

- 1K LOC⇒0.477
- 5K LOC⇒0.511
- 10K LOC⇒**0.795**

■ 1K LOC と5K LOC より大きい開発データでは 全データと相関係数に差異はなかった

■ 10K LOCより大きい開発データでは、上工程の設計製造工数で0.7以上と比較的強い相関がSTバグ数と見られた(表では特に相関が強かった5つの項目のみ表示)

項目名	相関係数
上工程設計製造工数	0.725
上工程設計製造工数(CD除く)	0.675
FD設計製造工数	0.666
DD設計製造工数	0.795
CD設計製造工数	0.759

■ **10K LOCより大きい開発データを対象**にモデル式を作成する

6. 重回帰分析によるモデル式の作成

説明変数の選択

STバグ数を目的変数として各項目間で重回帰分析を実施

今回説明変数は以下の2つを選択

- 上工程設計製造工数
- 上工程レビュー工数

説明変数の選択理由

- 上工程設計製造工数
 - STバグ数と相関係数が特に強かった(0.725)
 - STバグの作り込み工程を調べてみるとFD 29%、DD 24%、CD 47%と各工程にばらついており、上工程全体の項目を用いた方が妥当
- 上工程レビュー工数
 - 上工程設計製造工数を説明変数として選択した場合に、他の項目よりも調整済決定係数が高くなる
 - モデル式に適用した場合に負の値を取る(開発の実体と合っている)

重回帰分析実施結果

重回帰分析を実施した結果

	目的変数名	残差平方和	重相関係数	寄与率R^2	R*^2
	STバグ摘出件数実績	191.904	0.747	0.558	0.530
		R**^2	残差自由度	残差標準偏差	
		0.503	31	2.488	
vNo	説明変数名	残差平方和	変化量	分散比	偏回帰係数
0	定数項	196.486	4.581	0.7400	0.938
5	開発規模(実績)	189.099	-2.806	0.4451	-
6	自製規模(実績)	186.365	-5.539	0.8917	-
14	上工程バグ摘出件数実績	182.366	-9.538	1.5691	-
17	全体工数実績(ST除く)	185.194	-6.710	1.0870	-
18	テスト工数実績(ST除く)	185.194	-6.710	1.0870	-
20	上工程設計製造工数実績	434.196	242.292	39.1395	0.001
21	上工程レビュー工数実績	206.339	14.435	2.3318	-0.002

重相関係数で0.747、決定係数で0.558、調整済決定係数で0.530とある程度の精度の高い値が得られた

モデル式の作成

重回帰分析を実施した結果から得られた値を使用してモデル式を作成

$$Y = 0.938 + (0.001 * X_1) - (0.002 * X_2)$$

Y : ST予測バグ数

X1: 上工程設計製造工数実績

X2: 上工程レビュー工数実績

7. 作成したモデル式の評価

モデル式の評価

■ 作成したモデル式を見ると、STバグ数に対して上工程設計製造工数が正の(STバグ数が増える)要素を与え、上工程レビュー工数が負の(STバグ数を減らす)要素を与えている

- 設計製造を実施すればバグが埋め込まれ、レビューでそのバグを排除するという開発の実体と整合が取れており、今回作成したモデルが妥当であることを示している

モデル式

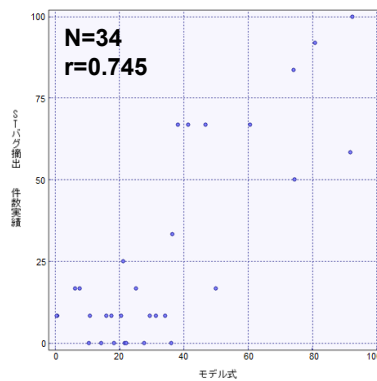
$$Y = 0.938 + (0.001 * X_1) - (0.002 * X_2)$$

赤色箇所: 上工程設計製造工数の要素

青色箇所: 上工程レビュー工数の要素

モデル式の評価

■ 自製規模が10K LOCより大きい開発のデータを対象に作成したモデル式によって得られたST予測バグ数とST実績バグ数との散布図と相関係数を出した



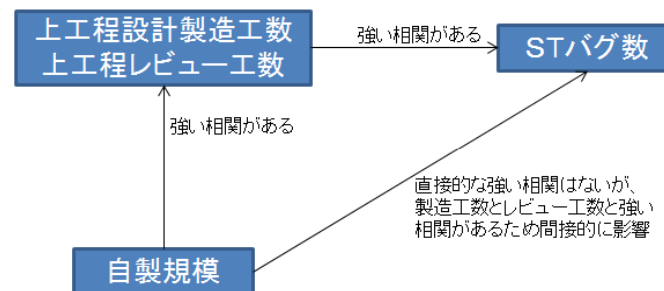
■ 相関係数が0.745と高い値が得られたことから、今回作成したモデル式は自製規模が10K LOCより大きい開発に対しては、充分適用可能な精度だと判断

8. 考察

考察

自製規模は直接的にはSTバグ数と強い相関は見られなかったが、
間接的にSTバグ数に影響を及ぼしていると考えられる

- モデル式で説明変数とした上工程設計製造工数と上工程レビュー工数は自製規模と強い相関(相関係数0.81程度)がある
⇒ 自製規模が大きくなるほど上工程設計製造工数と上工程レビュー工数も多くなるため、自製規模は間接的にSTバグ数に影響を及ぼしている



考察

上工程設計製造工数と上工程レビュー工数の2つの説明変数が、一般的にバグ数と相関が強いとされる自製規模より相関が強くなった要因

- 同じ自製規模の開発でも、一般的に難易度の高い開発ほど上工程設計製造工数と全体バグ数が多くなることが想定される
⇒ 全体バグ数が多くなことに比例して、ST工程で検出されるバグ数も多くなることが考えられる(全体バグ数とSTバグ数の相関係数は0.432)

開発難易度によって値に変動が出てくる上工程設計製造工数の方が自製規模よりSTバグ数と強い相関が出たことが考えられる

- 同じ設計製造工数をかけた成果物でも、レビュー工数をかけた方が、よりバグを排除できる可能性が高くなり、結果的にSTバグ数の削減につながる

バグを作り込む設計製造工数とバグを排除するレビュー工数は密接な関係があり、上工程設計製造工数単体より、上工程レビュー工数を説明変数として組み合わせることで、よりSTバグ数に対する相関が強くなったと考えられる

考察

モデル式で説明変数として選択した上工程設計製造工数と上工程レビュー工数に対する標準偏回帰係数を算出した

目的変数名	残差平方和	重相関係数	寄与率 R^2	R^{**2}	残差自由度	残差標準偏差
STバグ摘出件数実績	191.904	0.747	0.558	0.503	31	2.488
変数名	偏回帰係数	標準誤差	t 値	標準偏回帰	トレランス	
定数項	0.938	1.080	0.860			
上工程設計製造工数実績	0.001	0.000	6.256	0.763	0.959	
上工程レビュー工数実績	-0.002	0.001	-1.527	-0.186	0.959	

- 標準偏回帰係数より、**上工程設計製造工数**の方がSTバグ数に与える影響度が大きいことを統計的に示していると言える

今回作成したモデル式で使用している説明変数はテスト工程での測定項目がなく、**全て上工程完了時点には実績値が得られる項目**となっている

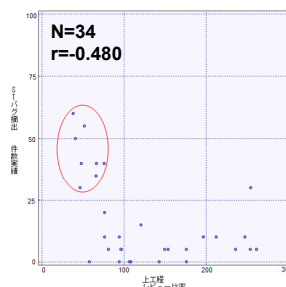
- 開発の早い段階で開発状況の異常(システムテストでバグが多発する要因)を検知するという目的を満たしている点も評価できるポイント

考察

モデル式を見るとSTバグ数に対して設計製造工数が正の要素、レビュー工数が負の要素を与えている

- 上工程全体工数(設計製造工数+レビュー工数)に対する**上工程レビュー工数の比率**(**上工程レビュー工数 / 上工程全体工数**)がSTバグ数を削減するため重要な要素になる

自製規模が10K LOCより大きい開発のデータを対象に上工程レビュー比率とSTバグ数の散布図を出してみた



上工程レビュー比率のある値を100とした場合、100を切っている開発でSTバグが多発(図の赤丸箇所)していることが分かる

モデル式と併用して、**上工程レビュー比率も1つの基準値として設けるべきである**

モデル式と基準値の使用方法

■ 基準値の使用方法(上工程レビュー比率)

- 上工程完了時点で上工程レビュー比率が低くなる状況を未然に防止するため、**開発開始時**に適用する

■ モデル式の使用方法

- **上工程完了時**に適用する
- STバグ数が多発するような予測値になった場合
 - 開発側にレビュー不足や検討不足がないことを確認してもらう
 - 品質保証部門が開発成果物を確認する対象とする
(品質保証部門が全ての開発成果物を確認するのは工数面から見て不可能なため、STバグが多発する予測値になったものを重点的に確認する)

9. 今回の成果と今後の課題

今回の成果と今後の課題

■ 本研究の成果

- 自製規模が10K LOCより大きい開発に対してSTバグ数の予測モデルを構築することができた
- 作成したモデルからSTバグ数に影響してくる要因をプロセスメトリクスから特定できた
(上工程レビュー比率が低いとSTバグが多発している開発が多い)

■ 今後の課題

- 自製規模が10K LOCより大きい開発しかモデルが作成できなかった
- 今回作成した予測モデルが実際の開発への適用実績がない
⇒ 今後、これらの課題を解決していくことで、STバグ予測モデルの精度を向上させ、ST工程でバグが多発することを未然に防止することで、出荷納期遅延のリスク低減を目指していきたい

NECグループビジョン2017

人と地球にやさしい情報社会を
イノベーションで実現する
グローバルリーディングカンパニー



Empowered by Innovation

NEC