

受託開発ソフトウェアの保守に おける作業効率の要因

S161053 前田欣耶

背景

- 近年、受託開発ソフトウェアの保守にかんして、ベンダとユーザで契約を結ぶ場合があり、保守の重要性は高まっている。

この論文ではユーザ企業のソフトウェア保守のコストを抑えることを目的とし、保守の作業効率と関連のある特性を明らかにしている。

目次

- 1 データセット
- 2 技術者あたり保守量の定義
- 3 分析
- 4 まとめ

1 データセット

・分析対象のデータセットは、財団法人経済調査会によって2007年度に83組織から収集されたケースである。(表1)

・一つのケースが複数のシステム構成を持つことがあるため、クライアントサーバ、web系、メインフレームの3つのダミー変数を作成して分析する。

表1 測定量一覧

測定量名	詳細
保守標準化	保守プロセスの標準化の状況 (標準化済=1, 標準化作業中/未標準化=0) この測定量のみ, 組織全体における状況を表す
システム構成	保守対象ソフトウェアが動作するシステムの構成 (クライアントサーバ, Web 系, メインフレーム)
母体プログラム本数	保守対象ソフトウェアの総プログラム本数
保守プログラム本数	保守を行ったプログラムの本数
技術者数	常駐しているベンダの保守技術者数
人的要因	保守プロジェクトあるいは保守組織の大きさと熟練度
問題要因	対象とする問題の型と重要度, 要求仕様の構成, 問題解の制約, 波及度合い
プロセス要因	要求仕様言語, 設計/プログラミング言語, 開発方法論
プロダクト要因	対象システムの信頼性, 規模, 効率, 制御構造, 複雑度
資源要因	対象とするハードウェア, 期間, 予算
ツール要因	ライブラリ, コンパイラ, テストツール, 保守ツール, リバースエンジニアリングツール

2 技術者あたり保守量の定義

- 投入された人的資源 → 技術者数
- 産出量 → 1 年間の保守プログラム本数

↓
として

- 技術者あたり保守量(保守プログラム本数/技術者数)
→ 作業効率

技術者あたり保守量の統計量を表 2 に
保守プログラム本数と技術者数との関係を図 1 に

表 2 技術者あたり保守量の統計量

項目名		詳細
保守標準化		保守プロセスの標準化の状況 (標準化済=1, 標準化作業中/未標準化=0) この測定量のみ, 組織全体における状況を表す
システム構成		保守対象のソフトウェアが動作するシステムの構成 (クライアントサーバ, Web 系, メインフレーム)
母体プログラム本数		保守対象ソフトウェアの総プログラム本数
保守プログラム本数		保守を行ったプログラムの本数
技術者数		常駐しているベンダのソフトウェア保守技術者数
生産性変動要因	人的要因	保守プロジェクトあるいは保守組織の大きさと熟練度
	問題要因	対象とする問題の型と重要度, 要求仕様の構成, 問題解の制約, 波及度合い
	プロセス要因	要求仕様言語, 設計/プログラミング言語, 開発方法論
	プロダクト要因	対象システムの信頼性, 規模, 効率, 制御構造, 複雑度
	資源要因	対象とするハードウェア, 期間, 予算
	ツール要因	ライブラリ, コンパイラ, テストツール, 保守ツール, リバースエンジニアリングツール

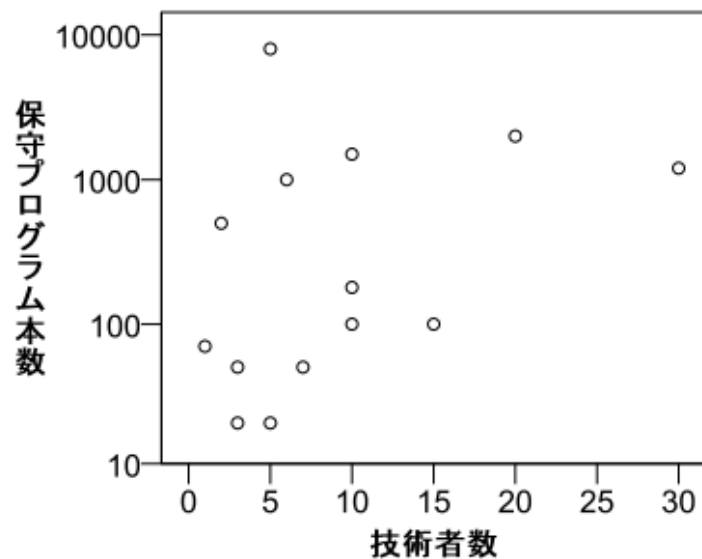


図 1 技術者数と保守プログラ本数との関連

3 分析

- 技術者あたり保守量を目的変数として重回帰分析を行った。

重回帰分析を用いたのは、各説明変数が目的変数に対してどの程度関連を持っているか個別に知るため

3.1 重回帰分析の適用方法

- 標準化状況、母体プログラム本数、システム構成、生産性変動要因を説明変数の候補とした
- 外れ値の影響をさけるために対数変換を適用
- 欠損地処理をするためにリストワイズ除去法の適用
- 変数増減法では回帰係数の検定のためのF値に対応するp値
0.05から0.5の範囲で0.2より小さければ変数を取り込み0.3より大きければ除外する

3.2 重回帰分析の妥当性

- 構築されたモデルの条件指標(表3)
- 各説明変数のVIF(表4)

表3 重回帰モデルの概要

決定係数	自由度調整済み決定係数	p 値	条件指標
0.78	0.70	0.01	26.54

・多重共線性

条件指標が30を超える場合発生

表4 採用された説明変数

測定量名	標準化偏回帰係数	p 値	VIF
母体プログラム本数	0.89	0.01	2.98
標準化状況	0.58	0.01	1.15
問題要因	-0.52	0.11	3.03

VIFが10を超える場合発生

3.3 重回帰分析の結果

表 3 重回帰モデルの概要

決定係数	自由度調整済み決定係数	p 値	条件指標
0.78	0.70	0.01	26.54

説明変数の説明力は比較的高い

表 4 採用された説明変数

測定量名	標準化偏回帰係数	p 値	VIF
母体プログラム本数	0.89	0.01	2.98
標準化状況	0.58	0.01	1.15
問題要因	-0.52	0.11	3.03

ユーザが制御しうる測定量は標準化状況のみ。

3.4 標準化の効果量

- プロセスの標準化が技術者あたり保守量にどの程度の影響を与えるか推定する。
- 構築されたモデルは以下のように表される。

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$$

y: 対数変換した技術者あたり保守量

x₁: 対数変換した母体プログラム本数

x₂: 標準化状況 x₃: 問題要因 β₀: 回帰定数

β₁~β₃: 偏回帰係数 ε: 誤差項

- 対数変換しない技術者あたり保守量 y' とすると
$$y' = \exp(\beta_0) \exp(\beta_1 x_1) \exp(\beta_2 x_2) \exp(\beta_3 x_3) \exp(\varepsilon)$$

よって、標準化されている場合、技術者あたり保守量に $\exp(\beta_2)$ を乗じる効果があることになる

表 5 技術者あたり保守量への効果

	点推定	下側信頼限界	上側信頼限界
偏回帰係数 (β_2)	2.08	0.62	3.55
効果 ($\exp(\beta_2)$)	8.04	1.86	34.69

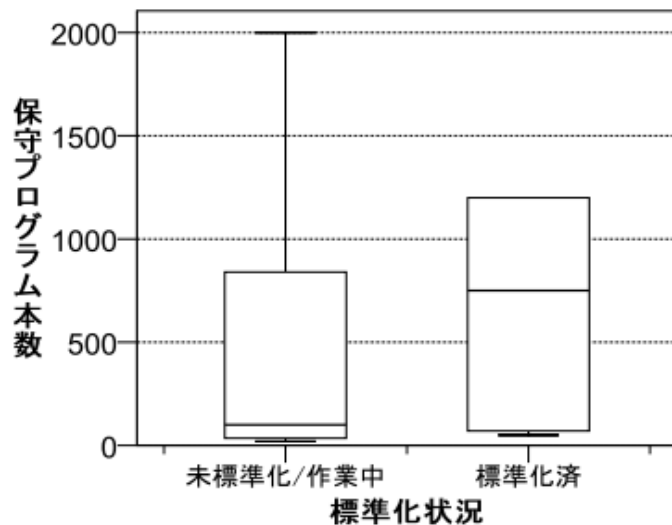


図 2 標準化状況と保守プログラム本数との関連

標準化済のほうが中央値が大きい
が大きく違う。

差も有意ではない。

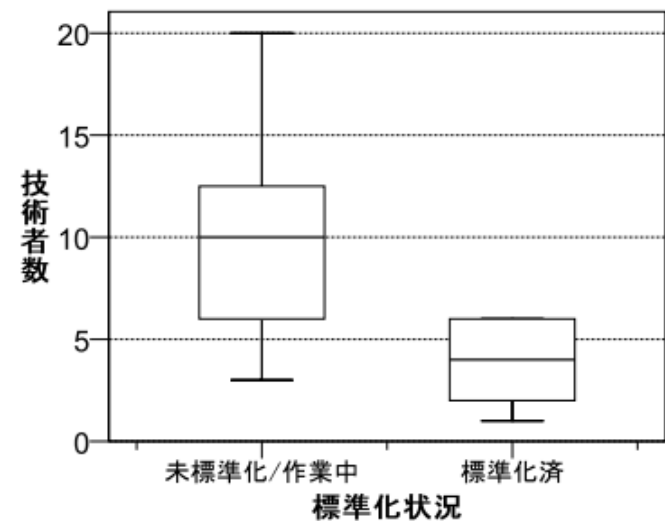


図 3 標準化状況と技術者数との関連

標準化済のほうが中央値が小さく
違いは見える。

だが差は有意ではない。

3.5 技術者を考慮した分析

- 技術者と標準化状況の関連は有意ではなかったが、図3を見ると関連があるように見える。
- 技術者数を説明変数に追加し同じ方法で重回帰分析を行った。
が、変化はなかったので、技術者数を考慮すると、技術者あたり保守量と標準化状況は関連があると言える。

4 まとめ

- ソフトウェア保守の作業効率に関連を持つ特性を明らかにすることが目的だった
- 技術者あたり保守量を作業効率と定義し、これを目的変数として重回帰分析を行った
- 結果保守プロセスの標準化状況と技術者あたり保守量とに関連があることがわかった
- プロセスを標準化することにより、技術者あたり保守量がだいたい8倍改善される。