

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 2 日現在

機関番号：34419

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25330090

研究課題名(和文) ソフトウェア保守・運用の効率化に向けた計画立案支援

研究課題名(英文) Planning support for efficient software maintenance and operation

研究代表者

角田 雅照 (TSUNODA, Masateru)

近畿大学・理工学部・講師

研究者番号：60457140

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：ソフトウェア保守・運用の効率化に向けた計画立案を支援するために、以下を明らかにした。(1) 作業時間に基づくソフトウェア保守ベンチマーキング、(2) 作業時間と単価に基づくシステム運用のベンチマーキング。また、ソフトウェア保守効率やシステム運用作業効率の予測モデル構築に向け、以下について明らかにした。(3) ソフトウェア開発プロジェクト予測モデルにおける外れ値の影響、(4) ソフトウェア開発プロジェクト予測におけるトービットモデルの効果。これらの成果を用いることにより、ソフトウェア保守、システム運用それぞれについて、最適な計画の立案が可能となり、その結果、それぞれの効率が高まることが期待される。

研究成果の概要(英文)：The goal of the research is to support making the plan of software maintenance and system operation. To support that, we clarified the followings, (1) benchmark of software maintenance based on working time, (2) benchmark of information system operation based on working time and unit cost. Also, to support building prediction models of maintenance efficiency and operation efficiency, we clarified the followings, (1) influence of outliers on software development project prediction, (2) advantages of Tobit model on software development project prediction. Using the benchmarks and analysis results, we can make appropriate maintenance plan and operation plan, and it will enhance efficiency of software maintenance and system operation.

研究分野：総合領域

キーワード：ソフトウェア保守 作業効率 コスト ベンチマーク 統計的分析 企業横断データ

1. 研究開始当初の背景

近年、ソフトウェアは社会のあらゆる場面で利用されており、それに伴い、情報サービス・ソフトウェアの市場規模も非常に大きくなっている。ただし、ソフトウェアのライフサイクルは開発だけで完結していない。ソフトウェアの開発後には、ソフトウェア保守とシステム運用が必要となる。ソフトウェアの保守とは、ソフトウェアの障害解析/復旧や性能強化/改善などを行うことであり、システム運用とは、システムの起動/終了や監視などを行うことである。これらソフトウェア保守とシステム運用についても、市場規模が非常に大きくなっている。矢野経済研究所の調査によると、2009年度のソフトウェア保守、ハードウェア保守、システム運用をまとめた市場規模は1兆1,422億円である。従って、ソフトウェア保守とシステム運用の効率を高める(できるだけ少ない投入コストで、ソフトウェアの保守、運用を行う)ことは非常に重要な課題となりつつある。

ソフトウェア開発については、プロジェクトの計画時に決める条件(要員数など)が、効率(生産性)にどう影響するかを分析した研究が多数存在しており、それらの定量的な分析結果に基づいて「ソフトウェア開発の効率を高めるためには、どのように開発計画を立案すべきか」という指針がある程度明らかとなっている。例えば研究代表者らの研究結果[5]によると、(開発期間に厳しい制約がなければ)開発期間を長めに設定するとともに、開発要員数を抑えることにより、効率を高めることができるといえる。

これに対し、ソフトウェア保守とシステム運用に関しては、保守・運用計画時に決められる条件(要員数など)の効率への影響を分析した研究は非常に少なく、定量的なデータ分析結果に基づいて「ソフトウェア保守とシステム運用の効率を高めるためには、どのように保守・運用計画を立案すべきか」が明確にされているとはいえないのが現状である。ソフトウェア保守の効率に影響する要因を分析した研究はいくつか存在するが、ソフトウェア保守の計画立案時に決められる条件(要員数など)と効率との関係を分析した研究は非常に少ない。それらの条件と効率との関係を定量的に分析・調査したものとして、経済産業省と日本情報システム・ユーザ協会がまとめた調査が挙げられるが、ソフトウェア保守の効率に影響する条件全てがカバーされているとはいえない。

システム運用に関しては、Information Technology Infrastructure Library (ITIL) というベストプラクティス集が存在するが、「ある条件(要員数など)を計画時に変更した場合、どの程度効率が変化するのか」といったことが定量的に示されているとはいえない。また、上述のように、保守・運用計画時に決められる条件(要員数など)の効率への影響を分析した研究は非常に少ない。

2. 研究の目的

本研究の目的は、ソフトウェアの保守とシステム運用それぞれの効率が高まるように、計画立案を支援することである。そのためのアプローチとして、企業横断的に(複数の企業から)収集したデータを用いて、それぞれの作業効率ベンチマークの作成を行った。また、当初は汎用的な効率予測モデルの作成を目指したが、用いたデータセットでは精度の高いモデルが構築できなかったため、各ユーザ、ベンダが精度の高い効率予測モデルを構築するために必要な要素技術を研究した。

3. 研究の方法

以下の手順により、ソフトウェア保守及びシステム運用の効率ベンチマークを作成した。

(1) 分析対象の変数の選定

以下の点に注意して、データセットから分析対象とする変数を選定する。効率は投入工数に対する保守実績(ソースコード修正量)と定義する。

- a. 目的変数(効率)に影響を持ちうる変数かどうか
- b. 効率ベンチマークと効率予測モデル利

(2) 外れ値の除去

(1)で選定した変数から、外れ値を除去する。外れ値とは、身長で例えると、170cmと入力するところを1700cmと入力されたようなデータである。

(3) 効率と各変数との関連の分析

ソフトウェア保守計画立案時に決定する条件のうち、効率に影響する(効率と関連の強い)ものを、(1)で選定した変数を用いて明らかにする。

(4) 効率ベンチマークの作成

(3)で明らかにした、効率の違いに影響する条件(要因)でデータを層別し、効率のベンチマークを作成する。

4. 研究成果

(1) 作業時間に基づくソフトウェア保守ベンチマーキング

ソフトウェア保守における作業効率のベンチマークを確立するために、企業横断的データを用いて、作業効率に影響する特性を明らかにした。予備分析では、保守費用と作業時間との関連が強いことを示すと同時に、実施FPの代わりに用いる修正量として、実施PG本数が最適であることを示した。この結果に基づき、作業効率を $\text{実施PG本数} \div \text{作業時間}$ と定義し、分析を行った。

分析では、保守のベンチマーキングのため

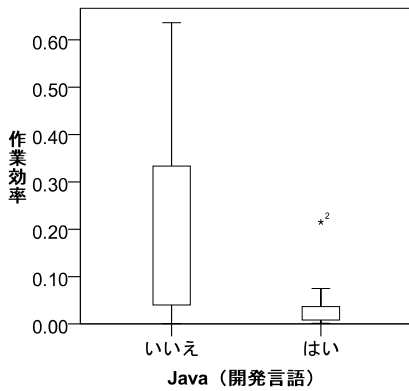


図 1 Java の作業効率

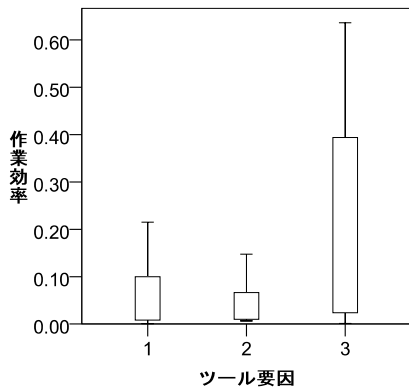


図 2 ツール要因の作業効率

に、特異別に作業効率の箱ひげ図（図 1，図 2）を示すとともに、開発言語、業種などの各特性が作業時間に与える影響を重回帰分析により確かめた。作業時間を増加させる特性は、作業効率を低下させる特性である。

修正 PG 本数と各特性を説明変数とし、作業時間を目的変数として重回帰分析した結果、Web 系（システム構成）、Java（開発言語）、ツール要因（生産性変動要因）、金融・保険業（業種）、適応保守（保守種別）が作業時間との関連があった。これらの特性を同時に説明変数として用いた場合、Java（開発言語）とツール要因（生産性変動要因）が作業時間に関連があった。

本研究の主要な貢献は、多数の企業が含まれた企業横断的データを用い、作業時間に基づいて定義した作業効率を用いて、作業効率に関連する特性を明らかにしたことである。

(2) 作業時間と単価に基づくシステム運用のベンチマーキング

委託側企業がシステム運用費用を見直す際などに、費用の妥当性判断の参考となるような情報の提供を目指し、システム運用費用に影響を与える要因の分析を行った。具体的には、作業時間と技術者の単価から簡易的に価格を推定することを前提とし、作業時間と

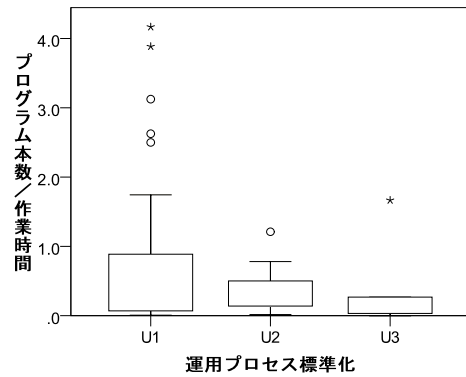


図 3 運用プロセス標準化と作業効率との関係

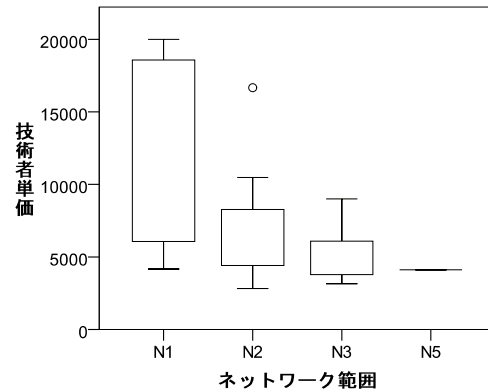


図 4 ネットワーク範囲と技術者単価との関係

単価に影響する要因をそれぞれ分析した。分析の結果、以下が明らかとなった。

- システム運用費用（契約金額）は、受託側作業時間によって大部分が決定している。
- 年間総作業時間は、プログラム本数と最大利用者数から決まるが、それらだけでは作業時間を高い精度で推定することはできない。
- 運用プロセスの標準化は作業効率を改善する可能性があるが、明確な傾向あるとまではいえない（図 3）。
- 社会的影響度は作業効率と関連を持つ可能性があるが、関連はあまり強くない。
- ネットワーク範囲が狭い場合、単価が低くなる傾向がある（図 4）。
- 契約形態が請負の場合、単価が低くなる可能性がある。
- SLA を締結していない場合、単価が低くなる可能性があるが、明確な傾向あるとまではいえない。

(3) ソフトウェア開発プロジェクト予測モデルにおける外れ値の影響

精度の高い効率予測モデルを構築するために、外れ値について分析を行った。分析では工数見積もりを取り上げて実験した。外れ値により工数見積もりの精度が低下すること

を防ぐため、外れ値除去法が適用される場合がある。外れ値除去法とは、見積もりの根拠データに含まれるプロジェクトが外れ値であるか検証し、外れ値と判断されたものをデータから除去する手法である。例えば、重回帰分析によりソフトウェア開発工数を見積もる場合、Cookの距離を用いて外れ値を除去すること多い。

本研究では、定量的にソフトウェア開発工数を見積もる場合に、データに含まれる外れ値がどの程度影響を与えるのかを分析した。そのために、本研究ではデータセットに外れ値を実験的に追加することを行った。具体的には、目的変数である工数と、説明変数として最も重要な開発工数に外れ値を含めた。実験では以下の2つのパラメータを変更した場合の見積もり精度の変化を確かめた。

- 外れ値が含まれる割合
- 外れ値の程度

さらに、データに外れ値が含まれているパターンについても変更して分析した。具体的には以下のパターンでデータに外れ値を含める。

- 見積もり根拠データ：外れ値あり，見積もり対象データ：外れ値なし
- 見積もり根拠データ：外れ値あり，見積もり対象データ：外れ値あり
- 見積もり根拠データ：外れ値なし，見積もり対象データ：外れ値なし

表1に実験結果の一部を示す。実験の結果、開発規模に外れ値が含まれていても、外れ値の割合が10%で、外れ値の程度が100%の場合、あまり影響が大きくないことなどがわかった。また、類似性に基づく見積もり方法の場合、テストデータに外れ値が含まれていても、比較的影響が小さい可能性があることがわかった。

(4) ソフトウェア開発プロジェクト予測におけるトービットモデルの効果

精度の高い効率予測モデルを構築するために、トービットモデルを用いた場合の予測モデルの精度について評価した。分析では欠陥数予測を取り上げて実験した。トービットモデルは、目的変数の最小値が0であることを前提としたモデルである。このモデルを欠陥数予測に適用することより、適切な予測モデルが構築され、予測精度が改善することが期待される。実験において、実際のソフトウェア開発企業から収集されたデータセットを用いて、従来の重回帰分析とトービットモデルによる予測モデルを構築し、各モデルの予測精度を比較することにより、トービットモデルの有用性を評価した。各モデルの構築

表1 重回帰分析における、外れ値と見積もり精度の関係

フィットデータ	テストデータ	AE平均	AE中央値
外れ値あり	外れ値なし	-18.03	19.90
外れ値なし	外れ値あり	-110.67	-109.37
外れ値あり	外れ値あり	-107.91	-60.70

MRE平均	MRE中央値	BRE平均	BRE中央値
4.00%	-1.23%	-0.09%	-0.70%
-3.94%	-2.00%	-2.58%	-3.59%
1.28%	-2.40%	-1.97%	-2.17%

表2 各モデルの予測精度（対数変換適用時）

	AE平均値	AE中央値
Log-OLS	8.24	1.86
Log-Tobit	8.45	1.53
差分	-0.21	0.33

BRE平均値	BRE中央値	Pred25
226%	107%	21%
260%	93%	28%
-34%	14%	7%

前に、対数変換を各変数に適用した場合の予測精度についても確かめた。トービットモデル自体はかなり以前に提案されたモデルであるが、我々の知る限り、これまでソフトウェア欠陥数の予測に適用されておらず、その有効性は明らかでなかった。

対数変換を適用しない場合: OLS（最小二乗法に基づく重回帰モデル）とTobit（トービットモデル）を比較すると、5個中3個の評価指標の差分が正の値となっていた。特にBRE中央値が50%程度改善していた。ただし、BRE平均値とPred25の差分が負であった。BRE平均値についてはOLSよりも100%程度悪化していた。また、Pred25についても、値が8%であり、精度が高いとはいえなかった。このことから、対数変換を適用しない場合、予測精度を高めるために、最小二乗法に基づく重回帰モデルとトービットモデルの、どちらのモデルを適用すべきであると結論付けることは難しかった。

対数変換を適用した場合: 実験結果の一部を表2に示す。対数変換を適用した場合、対数変換を適用しない場合と比較すると、最小二乗法に基づく重回帰モデル、トービットモデルとも予測精度が大きく改善していた。Log-OLSとLog-Tobitを比較すると、5個中3個の評価指標の差分が正の値となっていた。AE平均値の差分は負の値であったが、差は非常に小さかった。Log-OLSと比べ、BRE平均値はLog-Tobitのほうが34%低かったが、対数変換しない場合と比較すると、両者の差

は小さくなっていた。BRE 中央値と Pred25 については、Log-OLS と比較してそれぞれ 14%、7%改善していた。これらの結果より、トービットモデルについても対数変換を適用することにより予測精度が改善するといえる。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 4 件)

1. A. Monden, M. Tsunoda, M. Barker, and K. Matsumoto, "Probing Software Engineering Beliefs about System Testing Defects: Analyzing Data for Future Directions," IEEE IT Professional, 印刷中, 査読有.
2. 角田雅照, 門田暁人, 松本健一, "組込みソフトウェア開発における設計関連メトリクスに基づく下流試験欠陥数の予測," SEC journal, Vol.11, No.2, pp.16-25, September 2015, 査読有.
3. 角田雅照, 戸田航史, 伏田享平, 亀井靖高, Meiyappan Nagappan, 鷗林尚靖, "上流工程での活動実績を用いたソフトウェア開発工数見積もり方法の定量的評価," コンピュータソフトウェア, Vol.31, No.2, pp.129-143, May 2014, 査読有.
4. 角田雅照, 門田暁人, Jacky Keung, 松本健一, "熟練者判断を取り入れたソフトウェア開発工数見積もりモデル," 情報処理学会論文誌, Vol.55, No.2, pp.994-1004, February 2014, 査読有.

[学会発表](計 10 件)

1. M. Tsunoda, A. Monden, K. Matsumoto, S. Ohiwa, and T. Oshino, "Analysis of Information System Operation Cost Based on Working Time and Unit Cost," In Proc. of International Conference on Computer and Information Science (ICIS 2016), June 28, 2016 (Okayama, Japan).
2. K. Ono, M. Tsunoda, A. Monden, and K. Matsumoto, "Influence of Outliers on Analogy Based Software Development Effort Estimation," In Proc. of International Conference on Computer and Information Science (ICIS 2016), June 28, 2016 (Okayama, Japan).
3. M. Tsunoda, A. Monden, K. Matsumoto, S. Ohiwa, and T. Oshino, "Benchmarking Software Maintenance Based on Working Time," In Proc. of Applied Computing and Information Technology (ACIT 2015), pp.21-28, July 15, 2015 (Okayama, Japan).
4. 村上優佳紗, 角田雅照, 戸田航史, "トービットモデルに基づくソフトウェア欠陥数予測の試み," ソフトウェアエンジニアリングシンポジウム 2015 論文集, 情報処理学会 SES2015, pp.77-82,

September 8, 2015 (東京).

5. 小野健一, 角田雅照, 門田暁人, 松本健一, "ソフトウェア開発工数見積もりにおける外れ値の実験的評価," ソフトウェア・シンポジウム 2015, pp.115-123, June 6, 2015 (和歌山).
6. 村上優佳紗, 角田雅照, 戸田航史, "ソフトウェア欠陥数予測におけるトービットモデルの適用," ソフトウェア・シンポジウム 2015, pp.172-174, June 6, 2015 (和歌山).
7. 小野健一, 角田雅照, "ソフトウェア開発工数見積もりにおける外れ値の影響評価," 情報処理学会研究報告, ソフトウェア工学研究会, 2015-SE-187, No.7, p.1-7, March 12, 2015 (東京).
8. 角田雅照, 門田暁人, 松本健一, 大岩佐和子, 押野智樹, "作業時間と単価に基づくシステム運用費用の分析," 情報処理学会研究報告, ソフトウェア工学研究会, 2014-SE-186, No.12, p.1-7, November 14, 2014 (大阪).
9. 角田雅照, 門田暁人, 松本健一, 大岩佐和子, 押野智樹, "作業時間に基づくソフトウェア保守ベンチマーキングの試み," 情報処理学会研究報告, ソフトウェア工学研究会, 2014-SE-185, No.3, p.1-7, July 9, 2014 (北海道).
10. 小野健一, 角田雅照, "企業横断的ソフトウェア保守データにおける個別企業の影響," 情報処理学会研究報告, ソフトウェア工学研究会, 2014-SE-184, No.7, p.1-7, May 20, 2014 (茨城).

[図書](計 0 件)

[産業財産権]

○出願状況(計 0 件)

○取得状況(計 0 件)

[その他]

ホームページ等

<http://www.info.kindai.ac.jp/~tsunoda/publication.php>

6. 研究組織

(1) 研究代表者

角田 雅照 (TSUNODA, Masateru)

近畿大学・理工学部・講師

研究者番号: 60457140