

システム基盤構築工数見積りモデルの継続的改善と普及展開

井ノ口 伸人^{1,a)} 大杉 直樹¹ 伏田 享平¹ 渡辺 絢子¹ 吉野 順¹
 藤貫 美佐¹ 渡辺 真太郎¹ 戸村 元久¹ 木谷 強¹

受付日 2014年5月16日, 採録日 2014年11月10日

概要: 本稿では, アプリケーションソフトウェアを稼働させるための土台であるシステム基盤構築の工数見積りモデルの継続的な改善と普及展開について述べる. アプリケーションソフトウェア開発の工数見積りモデルについては多くの研究で有効性が報告されているが, 企業の IT システム開発実務で広く利用されているとはいえない. システム基盤構築については工数見積りモデルの研究は少なく, 利用事例も少ない. 本稿では, NTT データにおけるシステム基盤構築工数見積りモデルの継続的な改善と利用範囲の拡大について事例を報告する. 規模を表すパラメータ 8 種類, 基盤構築の難易度や能力を表すパラメータ 12 種類を, プロジェクトマネージャからアンケート調査でデータを収集した. 重回帰分析で予測モデルを作成し, 社内, 国内グループ会社へ普及展開を行った. 2008 年度 2 つの部署のプロジェクト 10 件のデータから作成した見積りモデルを文書化して当該部署に提供していた取り組みは, 2012 年度 20 の部署から収集したデータ 43 件から作成した予測モデルを Web アプリケーションに組み込んで全国内グループ会社へ提供するに至った. 2008 年度に 0.31 であったモデルの相対誤差中央値は, 2012 年度に 0.34 に多少悪化した. 一方, 見積りモデル作成に利用したデータに基づき, 2008 年度はサーバ台数が 3~15 台の小規模システムのみを適用可能プロジェクトとしていたが, 2012 年度にサーバ台数 3~70 台の中大規模システムまで適用可能範囲を拡大した. 2008 年度 5 名であったモデルの年間利用者数は, 2012 年度には 479 名まで増加した.

キーワード: システム基盤構築, 工数見積り, 見積りモデル, 実証的ソフトウェア工学, ソフトウェア工学の実践

Continuous Improvement and Deployment of a Series of Effort Estimation Models for System Platform Development

NOBUTO INOGUCHI^{1,a)} NAOKI OHSUGI¹ KYOHEI FUSHIDA¹ JUNKO WATANABE¹ JUN YOSHINO¹
 MISA FUJINUKI¹ SHINTARO WATANABE¹ MOTOHISA TOMURA¹ TSUYOSHI KITANI¹

Received: May 16, 2014, Accepted: November 10, 2014

Abstract: This paper reports continuous improvement and deployment of a series of effort estimation models for system platform development. Many studies have reported the effectiveness of effort estimation models; however few of them have used in companies' practical system development. This paper reports an example of deployment of a series of estimation models in NTT DATA Corporation with continuous improvement. In order to derive the model, we first collected 8 metrics for sizing and 12 metrics for measuring difficulty and team capability by a questionnaire for 10 project managers. We then derived the statistical model by regression analysis with the collected data. Although median of relative error of the model was 0.31 in 2008, it had been a little degraded to 0.34 in 2012; on the other hand, its applicability had been largely expanded from small-scale system development consisting of 3 to 15 servers, to medium- or large-scale system development consisting of 3 to 70 servers. While we first only provided the documented model to 5 practitioners in 1 company section, now we had developed a web application with the estimation model. This web application was used by 5 users only in 2008; but it had been grown up to 479 users in 2012.

Keywords: system platform development, effort estimation, estimation model, empirical software engineering, software engineering in practice

¹ 株式会社 NTT データ技術開発本部
 Research and Development Headquarters, NTT DATA Corporation, Koto, Tokyo 135-8671, Japan
^{a)} inoguchin@nttdata.co.jp

1. はじめに

IT システムの開発に要する工数を見積もるため, 多くの研究が行われてきた [1], [3], [4], [7], [17], [31], [33], [39].

特に開発するシステムの規模や工数に影響を与える要因をデータとして収集し、重回帰分析などで工数を算出する予測モデル（工数見積りモデル）は、過度の楽観視による大幅な見積り誤りを防止できるなどの有効性が報告されている [6], [14], [19], [20], [24], [38]. また、開発担当者が経験のないツールやフレームワークを利用する場合など、不確実性が大きい状況では、専門家が行うよりも見積りの精度が高い（正確に見積もれる）ことが報告されている [15], [28], [32], [34].

一方、企業のシステム開発実務で工数見積りモデルが広く利用されているとはいいいにくい [2], [13], [23], [25], [26], [40]. 著者が所属する NTT データでも、FP 法 [1] を始めとする工数見積りの支援専門部隊が全社研究開発部門に存在していたものの、収集したデータの基本統計量をまとめた報告書を年次発行するにとどまっていた。大幅な見積り誤りによる赤字プロジェクト発生のために工数見積りの重要性は議論されるものの、本稿の取り組み以前に工数見積りモデルの実務利用を全社的に進めた事例はなかった。

大幅な見積り誤りを防止するため、NTT データでは 2008 年度よりシステム基盤構築の工数見積りモデルについて継続的な改善とプロジェクトへの普及展開を進めている [16]. システム基盤構築（または単に基盤構築）は、ハードウェア、OS、ミドルウェア、ネットワークを組み合わせてアプリケーションソフトウェアを稼働させるための土台を作る作業であり、IT システム開発に要する工数の大きな部分を占める作業である。図 1 に NTT データにおけるプロセス定義と、そのなかでの基盤構築の位置付けを記す。

本稿では、図中破線部の作業を基盤構築と呼び、それらの作業に要する工数の見積りモデルについて述べる。NTT データでは、開発工程にあたる概念をプロセス、各プロセスで実施する作業をタスクと呼んで手順を定義している。本稿の見積りモデルでは、図中破線部に含まれる基盤の詳細設計、基盤制御機能開発、環境構築、運用ツール開発、基盤のテスト、システムテスト（基盤部分）として定義されたタスクの工数を見積もる。図 1 に示すように、これらのタスクは要件定義の後、アプリケーション開発、既存システムからの移行準備、運用手順書や体制整備、システム

のオペレータの研修などのサービス提供準備と並行して実施することが一般的である。

NTT データでは多くの場合、これらのタスクは基盤グループや方式担当などと呼ばれる基盤構築専任要員のチームが実施する。特に大規模システム開発においては、業務グループ、または AP 担当などと呼ばれるアプリケーション開発要員のチームと区別されて編成されることが多い。本稿の取り組みで報告する見積りモデルは、基盤グループが作業ごとの工数を積み上げて見積もった工数と比較することで、その妥当性を検証し、大幅な見積り誤りを防止することを目的としている。

本稿では、工数見積りモデルを企業のシステム開発実務で利用している事例として、NTT データにおける基盤工数見積りモデルの継続的改善と普及展開について述べる。取り組みの内容と実務者からのフィードバックについて報告することで、見積りモデルの実用化を検討している企業の実務者、PMO など支援組織の担当者、企業と共同研究を行う学術機関の研究者に、取り組みの進め方や期待される成果などの参考情報を提供する。

また本稿では、作成した見積りモデルの精度についても検証する。基盤構築工数見積りの事例 [21], [22], [29], [37] はいくつか報告されているものの、見積り精度が十分検証された事例は少ない。本稿では、モデル作成データやモデルパラメータの変化に応じた見積り精度の期待値の変化を、モデル作成データを統計分析することで検証する。また、見積り精度の期待値を従来研究 [26] と比較し、見積りモデルが実用に耐える精度を達成できたかどうかを検証する。

以降、2 章で関連研究として、企業における工数見積りモデルの実用事例、基盤工数見積りの研究を紹介する。3 章で NTT データにおける取り組みを紹介する。4 章で実プロジェクトのデータを利用して見積りモデルの精度の期待値を分析し、年度ごとの取り組みに応じた精度変化を検証する。5 章で既存研究との精度比較で実用性を考察し、見積り誤りが大きかったプロジェクトとその原因も考察する。さらに、利用者からのフィードバックについて考察し、6 章で結論と今後の課題を述べる。

2. 関連研究

2.1 企業における工数見積りモデルの実用事例

Moløkken ら [26] は過去のサーベイ [2], [13], [23], [25], [40] に基づき、企業における工数見積りモデルの利用を調査した。調査対象の国や対象者の役割、調査方法が異なるため結果を容易に一般化するのは難しいものの、いずれの調査でも最も多く利用されたのは専門家による判断（expert judgment）であり、工数見積りモデルを利用したという回答は多くなかった（全調査対象の 13~26%であった）。この調査によると、多くの企業で見積りモデルが実用化されているとはいいいくいと考えられる。

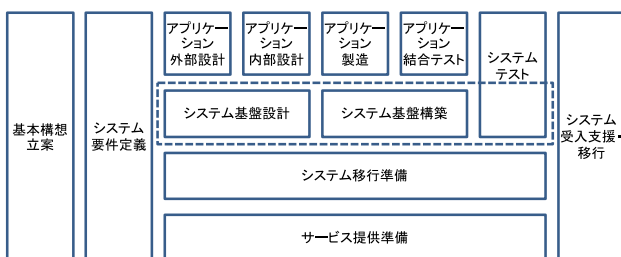


図 1 NTT データのプロセス定義におけるシステム基盤構築
 Fig. 1 Definition of system platform configuration in NTT data corporation's development processes.

表 1 日本のシステム開発企業における工数見積りモデルの利用事例
 Table 1 Practices of effort estimation model in Japanese system development companies.

企業名	見積対象	見積りモデルの概要	利用実績
富士通	Web システムが対象, 開発全体工数	Function Scale (簡易 FP) を規模パラメータとする COCOMO II [3] ベースモデル	2005 年度から適用を開始, 2006 年度時点で適用件数は 20 件~30 件.
日本ユニシス	汎用, 開発全体の工数	FP か SLOC を規模パラメータとする COCOMO II [3] ベースのモデル	2000 年より COCOMO II [3] を含めた複数手法による見積りを推奨. 利用実績は不明.
日本 IBM	汎用, 開発全体の工数	工数 = $\alpha \times \{10^{(\log \text{開発サイズ})} \beta\}^\gamma$ ※ α, β, γ は定数である.	開発した Web アプリケーションを社内ネットワークで提供. 12 年間運用し, 10,000 以上に適用.
ジャステック	汎用, 工程ごとのアプリケーション開発工数	設計書文字数か SLOC を規模パラメータとし, 規模 $\times (1 + \Sigma 13$ の変動要因) \times 生産性 $\times (1 + \Sigma 28$ の変動要因) で算出.	2003 年度, 同社における全プロジェクト (新規開発 150 件, 保守 95 件) に利用. 教育, 障害対応, 移行など (35 件) には未適用.
NTT データ ※本稿の内容	Web3 層構成システムが対象, 基盤構築工数	工数 = $\alpha \times$ 論理サーバ台数 $\beta \times$ 非機能要求グレード \times 基盤構築期間 γ ※ α, β, γ は定数, 詳細は 3 章で説明.	2008 年度から取組開始. Web アプリケーションを開発し, 社内ネットワークでグループ会社含めて提供. 2012 年度現在で利用者は 479 名.

※FP や SLOC などの規模を表すパラメータに, 規模あたりの開発工数 (文献 [9] [10] [11] では「生産性」と呼ばれている) を乗じるだけの算出方法は, 見積りモデルの利用事例ではないと判断した.

独立行政法人情報処理推進機構が日本のシステム開発企業における工数見積りモデルの利用事例について紹介している [9], [10], [11]. 文献 [9] では新規開発における見積りモデルの利用として表 1 の事例を紹介している. また, 文献 [10] では機能追加や維持保守開発, 文献 [11] ではテスト工程における見積りモデルの利用を紹介している. 日本 IBM やジャステックなど, ごく一部の企業では実務利用が進んでいる一方, 多くの企業で広く長期にわたって利用が進んでいるとはいえない.

表 1 の最下行に本稿で報告する NTT データにおける事例を併記する. 文献 [9] の事例と比較すると, 基盤構築工数の見積りモデルを実務利用している点が他と大きく異なる (文献 [10], [11] でも当該事例の紹介はない). また, 本稿の事例では, 文献 [9], [10], [11] では述べられていない見積りモデルの有効性についても検証する. 本稿の事例は稀少であり, 参考情報としての価値は高いと考えられる.

2.2 システム基盤工数見積りの研究

Valerdi [37] は, システム開発全体の工数見積りモデル COSYSMO (Constructive Systems Engineering Cost Model) を提案した. COSYSMO が見積もる工数には, 基盤の設計やシステムテストの作業も含まれるが, 基盤構築の大部分を占める基盤制御機能開発, 環境構築といったタスクは含まれない. また, COSYSMO の見積り対象は, システムの要件定義, マニュアル作成, 移行作業など, 基盤グループに割り当てられない可能性が高いタスクも含む. 本稿では, 基盤構築に要する作業の工数を必要十分に見積もるモデルの利用事例を報告する.

社団法人日本情報システム・ユーザ協会が日本の複数の企業から収集したシステム開発データを分析し, 開発の総工数, 総費用, 全体工期を互いに見積もるモデルを報告している [36]. 報告されている分析や見積りモデルは, システム開発全体の工数, 費用, 工期であり, 基盤構築のみが対象ではない. 本稿では, 基盤構築に要する作業の工数を必要十分に見積もるモデルの利用事例を報告する.

金子ら [21], [22] は, 大規模 IT インフラ構築工数の見積りモデルを提案した. Logical Partitioning, High Availability Cluster Multiprocessing, IBM WebSphere Application Server, IBM Universal Database といった基盤構成要素数を規模と考えて工数を見積もる. 本稿の見積りモデルは, これらの基盤構成要素数と同様の概念である論理サーバ台数に加え, システムに求められる非機能要求の高さを表す非機能要求グレード [12] と開発期間も変数として工数を見積もる. それらのパラメータを考慮することで, 見積り精度の改善が期待できる.

小澤 [29] は, サーバの種類と処理パターンから基盤構築工数を工程ごとに見積もるモデルを提案した. 典型的なサーバの種類と処理パターンについて実績データを分析し, 難易度係数を決定する. 構築する基盤のサーバの種類と処理パターンに対応する難易度係数から規模変数を算出し, 3 種類の線形モデルに代入して, 基盤設計・構築・テストの工数を各々見積もる. 本稿の見積りモデルも, 小澤の方法と同様に, 実績データを分析して係数を決定する. 加えて本稿では, 係数決定に利用した実績データを分析し, 小澤の研究では報告されていない見積りモデルの精度や実用性についても検証する.

3. NTT データにおける取り組み

図 2 に取り組みの手順を示す。2008 年度、基盤構築の大幅な見積り誤りの発生を契機として研究部門が取り組みを開始した。2008 年度に見積りモデルを作成し、文書化してデータ収集の協力者に配布した。2009 年度以降はモデルを改善し、Web アプリケーションに組み込んでサービスを提供した。図中の各手順では次を実施した。

(1) 対象プロジェクトの絞り込み

実用に耐える精度を得るために見積りモデルの対象とするプロジェクトを絞り込んだ。システム開発担当部署や基盤構築支援部署において基盤構築要員もしくはその管理者として勤め見積りの経験があり、見積り手法と結果の妥当性を客観的に評価できる能力を有する基盤構築の有識者と議論し、見積りモデルの想定利用範囲として下記のプロジェクトを対象とした。

- システム基盤を新規構築または完全更改

既存の基盤を拡張する場合より工数が多く必要であり、不確実性が高い。大幅な見積り誤りの可能性が高く、見積りモデルのニーズが大きいと思われる、基盤を新規構築または完全更改するプロジェクトを対象とした。

- システム基盤のアーキテクチャが Web 3 層構成

2008 年度時点で主流であったオープンシステムによる Web 3 層構成のアーキテクチャを対象とした。メインフレーム構成の基盤を新規構築することは稀であり、Hadoop などを利用する大規模クラスタ基盤、同様の基盤構成を仮想化して大量複製するクラウド基盤も、当時は稀であったため対象外とした。

- システム基盤を既存ミドルウェアの組み合わせで構築

性能要件が非常に厳しい場合など、ミドルウェアを開発・拡張する場合は対象外とした。基盤構築作業の一環として基盤制御のためのソフトウェアを開発・改良する場合、その工数は本モデルとは別に計算する必要がある。

(2) データ項目の選定とデータ収集

基盤構築に携わる有識者 7 名と議論し、基盤構築に影響を与えるパラメータを洗い出した。洗い出したパラメータ

を基盤の規模を表すもの、構築の難易度を表すもの、開発能力を表すものに分類し、影響の大きさを各有識者にアンケート調査した。影響が大きいと回答された表 2 の 20 パラメータを、収集するデータ項目として選定した。選定した項目について、2つの部署のプロジェクト 10 件からアンケート調査でデータを収集した。

(3) パラメータ決定、モデル作成

収集したデータを分析してパラメータを決定し、モデルを作成した。収集したデータはプロジェクト 10 件であったため、式 (1) に示すパラメータ 2 つの重回帰モデルを作成した。Cohen の効果量 [5] などに基づく単回帰モデルを作成することが妥当である可能性もあったが、有識者との議論で基盤構築の難易度や構築能力も含めたいという要望が強く、統計的妥当性の低下を許容して規模以外に 1 つの補正パラメータを含めた。

$$\text{工数} = \alpha \times \text{規模パラメータ}^\beta \times \text{補正パラメータ} \quad (1)$$

係数 α はシステム基盤の構築効率に相当する。組織に属する開発要員の能力や、開発対象基盤の難易度に依存し、この値が小さいほど、システム基盤構築に要する工数が小さくなる。係数 β は規模の増加が工数に与える影響の大きさを表す。 $\beta = 1.0$ であれば、規模と工数との関係は線形となり、規模の増加が工数に与える影響は一定である。 $\beta < 1.0$ であれば、規模の増加に比べ工数は緩やかに増加し、 $\beta > 1.0$ であれば、規模の増加に比べ工数が急激に増加する。

パラメータ決定では、回帰モデルで算出した予測工数と各プロジェクトの実績工数から、式 (2) で相対誤差 (MRE: Magnitude of Relative Error) を算出し、全プロジェクトの中央値 (MdMRE: Median of MRE)、MRE 0.25 以下で見積もれたプロジェクトの割合 (Pred25) を基準とした。MRE の値域は $[0.00, \infty)$ であり値が小さいほど精度が高いことを、Pred25 の値域は $[0.00, 1.00]$ で値が大きいほど精度が高いことを表す。

$$\text{MRE} = \frac{|\text{実績工数} - \text{予測工数}|}{\text{実績工数}} \quad (2)$$

また、各パラメータの分散拡大要因 (VIF: Variance Inflation Factor) を算出し、作成したモデルに多重共線性が発生していないことを確認した。多重共線性はモデルのパラメータどうしが強く相関している状態であり、見積り精度を大きく低下させる過適合の原因となる。VIF の値域は $[0.00, \infty)$ であり、5 を超える場合は多重共線性が発生している可能性が高い [8]。

まず規模パラメータの候補 (表 2 の項番 1~8) 各々について式 (1) の第 2 項、補正パラメータを除く単回帰モデルを作成し、精度が最も高くなる規模パラメータを選択した。次に、決定した規模パラメータの候補と補正パラメータの候補 (表 2 の項番 9~20) について式 (1) の重回帰モデル

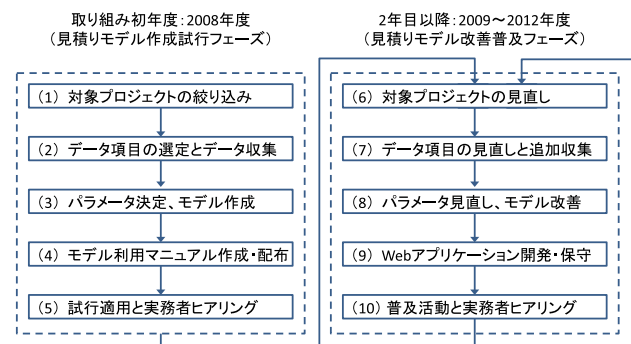


図 2 見積りモデル作成と普及改善の手順

Fig. 2 Procedure for deriving and improving estimation model.

表 2 選定された収集データ項目
Table 2 Selected data items for data collection.

分類	項番	データ項目名	説明	影響の大きさ※
規模	1	論理サーバ台数	システムを構成するサーバを、物理・仮想を区別せずに 1 台と数える。クラスタを構成するサーバは個別に数える。	9.00
	2	クラスタ構成数	システム内に含まれているクラスタの数。	12.00
	3	クライアント数	システムが管理するクライアントの数。	7.00
	4	ネットワーク機器数	上記以外の設計が必要な機器(ファイアウォールなど)の数。	8.00
	5	外部接続先数	システムが接続する外部システムの数。	4.00
	6	分散拠点数	地理的に分散している拠点の数。	6.00
	7	パッケージプログラム数	基盤として利用しているパッケージプログラムの数。	3.00
	8	テーブル数	管理しているデータベース中のテーブルの数。	6.00
難易度	9	許容計画停止頻度	運用上、許容される計画停止の頻度。	9.00
	10	稼働率	運転時間に対する実際の稼働時間。	13.00
	11	サービス中断時間	システム切り替えによる中断時間。	6.00
	12	RPO(目標復旧時点), RTO(目標復旧時間)	障害発生時、データを復旧すべき時点。障害発生時、データの復旧に要する時間。	6.00
	13	ピーク時レスポンスタイム, 多重度, TPS	ピーク時のレスポンスタイム, 同時アクセス数, 1 秒あたりに処理すべきトランザクション処理数。	6.00
	14	縮退時レスポンスタイム, 多重度, TPS	縮退時のレスポンスタイム, 同時アクセス数, 1 秒あたりに処理すべきトランザクション処理数。	11.00
	15	帳票スループット性能	帳票の単位時間あたりの出力枚数。	10.00
	16	情報の機密性	システムが取り扱う情報の機密密度。	6.00
	17	運用の自動化レベル	システム運用の自動化の度合い。	8.50
	18	許容移行停止時間	許容される移行のためのシステム停止時間。	12.00
	19	移行データ量・複雑さ	移行データの量および移行データの複雑さ。	6.00
開発能力	20	アーキテクチャの経験有無	類似アーキテクチャの構築経験有識者が体制にいるか否か。	1.00

※「影響の大きさ」に基盤構築工数への影響の大きさを示す。洗い出した全パラメータについて、7名の有識者に相対順位をアンケート調査した。1が最も影響が大きく、数字が小さいほど影響が大きい。7名の回答の平均値を上表に示す。

を作成し、精度が最も高くなる補正パラメータを選択した。2008年度に作成したモデルを式(3)に示す。規模パラメータとして論理サーバ台数(表2の項番1)、補正パラメータとして構築対象アーキテクチャの経験有無(表2の項番20)を選択した。式中、 α , β , 経験有無は係数を表す。機密情報のため α は非公開とし、他の係数、および各係数のVIFを表3に示す。作成したモデルのMdMREは0.31, Pred25は40%であった。VIFはいずれも1.00であり、多重共線性発生の可能性は低いと考えられる。なお、アーキテクチャ経験有の場合に係数が1.0となるようダミー変数化したため、VIFは算出していない。

$$\text{工数} = \alpha \times \text{論理サーバ台数}^\beta \times \text{アーキテクチャ経験} \quad (3)$$

表 3 見積りモデルの係数
Table 3 Coefficients of effort estimation model.

係数		2008	2009-2011	2012
β		0.92 (1.00)	0.98 (1.14)	0.75 (1.43)
γ		—	—	0.80 (1.56)
アーキテク チャ経験	有	1.0	—	—
	無	4.37 (1.00)	—	—
非機能要求 グレード	低	—	1.0	1.0
	中	—	1.60 (1.29)	1.91 (1.22)
	高	—	5.25 (1.37)	3.99 (1.45)

※()内に各係数に対するVIFの値を示す。また、—は当該年度のモデルにその係数が存在しないことを表す。

表 4 適用可能プロジェクトとプロジェクトデータ件数
Table 4 Applicable projects and # of projects in the data.

	2008	2009-2011	2012
構築種別	基盤を新規構築または完全更改		
アーキテクチャ	Web3 層構成		
構築方法	既存ミドルウェアの組み合わせ		
規模	3~15 台	3~30 台	3~70 台
構築期間	—	—	6~30 カ月
要員数	—	—	2~30 名
データ件数	10 件	33 件	43 件

※—は当該年度のモデルでは未定義であることを表す。

(4) モデル利用マニュアル作成・配布

作成した見積りモデルについて、モデルの利用方法、意味、作成方法や分析結果をモデル利用マニュアルとして文書化して配布した。データ収集に協力してくれた2つの部署10のプロジェクトに電子メール送付するとともに、当該部署の幹部と管理職に説明会を実施した。

(5) 試行適用と実務者ヒアリング

データ収集に協力してくれた部署の実務者5名に作成したモデルを試行的に利用してもらい、ヒアリングでモデルの改善方法と普及展開方法を検討した。モデルのさらなる改善や、継続的な改善の必要性を示唆する意見も多かったが、それらも含め、取り組みについて肯定的な意見が大半であった。得たフィードバックの詳細は5.3節に示す。

(6) 対象プロジェクトの見直し

データの分析結果および実務者から得たフィードバックに基づき、想定する見積り対象を見直した。より多くのプロジェクトに利用してもらえるよう、表4のように、見直した結果を「適用可能プロジェクト」として見積りモデルのマニュアルに明記した。

(7) データ項目の見直しと追加収集

表2に示すデータ項目を、見積り精度への寄与、利用時の入力のしやすさなどに基づいて見直した。有識者との議論に基づき、Putnamモデル[31]にならい、基盤構築期間を追加収集することとした。また、2008年度当初は表2の項目9~19の項目で基盤構築の難易度を収集していたが、項目どうしの相関が高く、多重共線性の原因となるため、非機能要求グレード[12]という1つの項目に統合して次の3段階で収集した。

- 低：システムトラブルの影響が特定の組織や企業内に閉じており、社会や業界まで広がる可能性が低いシステム。公共、法人分野の非基幹系システムを想定。
- 中：システムトラブルの影響が社会、または特定業界の一部に広がり、悪影響を及ぼすシステム。地方自治体のシステムや公共法人分野の基幹システムを想定。
- 高：システムトラブルが社会または特定業界全体に甚大な悪影響を及ぼすシステム。金融分野のミッションクリ

ティカルシステムや公共分野の中央省庁システムを想定。これらのデータ項目を、年間5~10件程度のプロジェクトについて継続的にヒアリングで収集した。表4に示すように、2008年当初2つの部署のプロジェクト10件であったデータは、2012年度には20の部署のプロジェクト43件に増加した。

(8) パラメータ見直し、モデル改善

データの収集状況、有識者や利用者との議論に基づき、パラメータを見直し、モデルを修正した。2009~2011年度に収集したデータを追加して修正したモデルを式(4)、さらに2012年度に収集したデータを追加して修正したモデルを式(5)に示す。

$$\text{工数} = \alpha \times \text{論理サーバ台数}^\beta \times \text{非機能要求グレード} \quad (4)$$

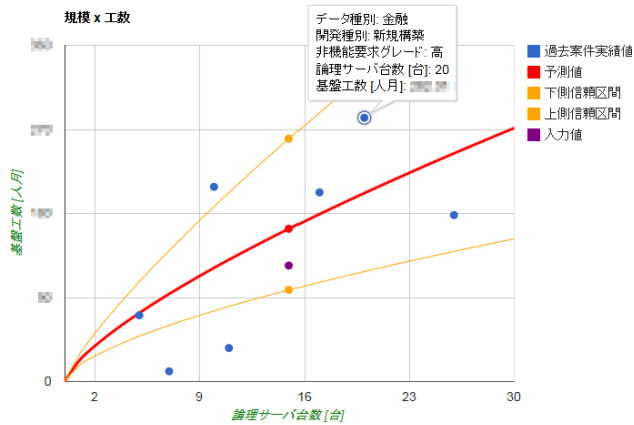
$$\text{工数} = \alpha \times \text{論理サーバ台数}^\beta \times \text{非機能要求グレード} \times \text{開発期間}^\gamma \quad (5)$$

式(4)、(5)では、主観的に判断される「アーキテクチャ経験」より、客観的に判断できるパラメータの方が、利用しやすいという利用者の意見を受け、補正パラメータを「非機能要求グレード」に変更した。「類似アーキテクチャの構築経験有識者が体制にいるか否か」について、開発要員のうち何人がどの程度の経験があれば「有識者がいる」とするかの判断が、判断する者の主観で変わりやすいため、使いにくいという意見が利用者からあがった。見積りモデルの利用者増加にともない、この問題の悪影響が大きくなることが予想されたため、より客観的に判断が可能な非機能要求グレードにパラメータを変更した。また、式(5)ではさらに「開発期間」を追加した。

機密情報のためαは非公開とし、他の係数、および各係数のVIFを表3に示す。2009~2011年度モデルのMdmREは0.46、Pred25は33%であった。また、2012年度モデルのMdmREは0.34、Pred25は44%であった。VIFは最大でも1.56であり、多重共線性発生の可能性は低いと考えられる。なお、非機能要求グレード低の場合に係数が1.0となるようダミー変数化したため、VIFは算出していない。

(9) Webアプリケーション開発・保守

2009年度に見積りモデルを組み込んだWebアプリケーションを開発した。アプリケーションの見積り結果表示画面例を図3に示す。横軸に論理サーバ台数(台)、縦軸に基盤構築工数(人月)をとる散布図上に、入力パラメータに対する見積り結果と見積り結果の95%信頼区間を折れ線グラフで表示する。また、実施予定プロジェクトの計画工数と、モデル作成に利用した過去プロジェクトを異なる色でプロット表示し、プロットにマウスポインタを重ねると図3のような情報を表示する。ユーザはプロジェクトの計画工数を、見積り結果を表す折れ線グラフや過去プロジェクトと比較することでその妥当性を検証できる。



※図中、縦軸の工数は機密情報のため、モザイク化した。

図 3 Web アプリケーションの見積り結果表示画面例

Fig. 3 An estimation result displayed on the Web application.

表 5 検証に利用したデータ

Table 5 Data used for the evaluation.

		2008	2009-2011	2012
データ件数		10 件	33 件	43 件
基盤構築工数 (人月) [比率尺度]	大規模	0	2	2
	中規模	0	4	5
	小規模	10	27	36
	変動係数	1.15	2.54	2.66
論理サーバ台数 (台) [比率尺度]	最大値	26.00	100.00	100.00
	中央値	8.00	10.00	10.00
	最小値	3.00	2.00	2.00
	変動係数	0.73	1.18	1.15
基盤構築期間 (月) [比率尺度]	最大値	23.00	28.00	28.00
	中央値	11.00	12.00	11.00
	最小値	4.00	4.00	4.00
	変動係数	0.44	0.50	0.50
アーキテクチャ経験有無 [名義尺度]	有/無			
非機能要求グレード [順序尺度]	低/中/高			

※ [] 内に尺度水準を記す。変動係数は標準偏差を平均値で除算した値であり、相対的なばらつきの程度を表す。機密情報のため最大値など記述できない基盤構築工数については、全データから外れ値の 1 件を除いて値域を 3 等分し、大規模 (600 人月以上)、中規模 (300 人月以上 600 人月未満)、小規模 (300 人月未満) の件数を示す。

(10) 普及展開と実務者ヒアリング

実施した普及展開活動の頻度とユーザ数の増加を表 6 に示す。2008~2009 年度当初は、主に数人の実務者や開発チームに対する個別説明、または部署の管理職や課長代理層十数名に対する説明会による普及展開を実施した。2010 年度からは、より大人数にアピールできる社内セミナーや社内展示会を企画・実施した。2012 年度には Web サイトに

表 6 実施した普及展開とユーザ数の増加

Table 6 Implemented promotions and growth number of users.

	2008	2009	2010	2011	2012
個別説明	8	12	3	0	3
利用説明会	2	2	3	0	0
社内セミナー	1	1	2	5	9
社内展示会	0	1	0	0	1
バージョンアップ	1	1	1	0	6
バージョンアップ 周知	0	1	0	0	2
読み物記事	0	0	0	0	3
年間ユーザ数※	5	48	102	193	479

※2008 年度は試用利用した実務者 5 名を、2009 年度以降は Web アプリケーションの月次ユニークユーザの年度計を記す。

読み物記事を掲載し、その読み物記事や公開している Web アプリケーションへの誘導記事を社内のポータルサイトに掲載することで、ユーザ数の大幅増加に成功した。2008 年度 5 名だったモデルの年間利用者数は、2012 年度には合計 479 名まで増加した。

普及展開の過程で実務者に見積りモデルや Web アプリケーションについてヒアリングを行い、適宜フィードバックを収集した。収集したフィードバックに基づいて、取り組みや普及展開の方針を決定・変更するとともに、継続的に Web アプリケーションを改善した。Web アプリケーションに大きな機能追加を行った場合は、社内広報システムを通じて国内全グループ会社に周知し、新規ユーザへの宣伝活動を兼ねてバージョンアップを通知した。実務者から収集したフィードバックの詳細を 5.3 節に、バージョンアップとバージョンアップ周知の頻度を表 6 に示す。社内セミナーや読み物記事の掲載に加え、これらの社内周知もユーザ数の向上に寄与したと考えられる。

実務者ヒアリングにおいて、複数の実務者から得られた典型的なフィードバックを次に示す。

- FP 法 [1] などアプリケーション開発の見積り方法は多く存在するが、基盤側の見積り方法は存在しないため有用である。
- 基盤構築工数の妥当性を顧客に説明する際に利用できた。
- 作業を積み上げて見積もった工数と、モデルの見積りが大きく変わらないことに驚いた。見積りの妥当性検証に利用できた。
- 設計共通化や再利用による工数削減 (生産性向上) 効果を、モデルの見積り結果と比較することで検証できた。
- 論理サーバ台数だけでなく、他の構成要素 (ネットワークスイッチ数など) も考慮すれば精度が向上すると考える。
- 基盤の仮想化、再利用、自動化ツール利用により、基

盤構築の生産性が近年大きく向上している。データ収集と再分析を継続実施し、モデル改善を続けて欲しい。

4. 見積りモデルの精度変化の検証

4.1 概要

取り組みによる見積りモデルの精度変化について検証するため、モデル作成に利用した表5のデータを分析し、データおよびモデルパラメータを2008年度、2009～2011年度、2012年度の状況に各々変更してMdMREとPred25を計算した。データについては、表5に示すとおり各年度でモデル作成に利用した10件、33件、43件を学習データおよびテストデータとした。年度ごとに別プロジェクトを選定してデータとしたわけではなく、年度ごとに逐次累積的に追加収集したデータであり、後の年度はそれ以前の年度のデータを含んでいる。モデルパラメータは各年度に応じて式(3)～(5)のように変更した。データ追加収集による精度の変化と、パラメータ見直しによる精度変化を評価するため、表7、表8に示すようにパラメータを固定してデータを変化させた場合、データを固定してパラメータを変化させた場合、データとパラメータいずれをも変化させた場合とで計9回MdMREとPred25を算出して比較した。

表7 精度変化の検証結果 (MdMRE)

Table 7 Evaluation result of accuracy changes with MdMRE.

データ件数 モデルパラメータ		2008	2009-2011	2012
		10件	33件	43件
2008	式(3)	0.31	(0.08) 0.57	(0.65) 0.55
		(0.70)	(0.34) (0.15)	(0.06) (0.14)
2009-2011	式(4)	0.28	(0.36) 0.46	(0.55) 0.34
		(0.43)	(0.97) (0.02)	(0.28) (0.10)
2012	式(5)	0.27	(0.94) 0.25	(0.80) 0.34

※各セルの値の左、上、左上に記述した()内に、左、上、左上の各セルの値との差の検定結果 (p 値) を示す。p ≤ 0.10 の場合、危険率10%で有意差があると判断し、p 値を下線で強調する。

表8 精度変化の検証結果 (Pred25)

Table 8 Evaluation result of accuracy changes with Pred25.

データ件数 モデルパラメータ		2008	2009-2011	2012
		10件	33件	43件
2008	式(3)	40%	(0.21) 18%	(0.78) 23%
		(1.00)	(0.71) (0.26)	(0.13) (0.34)
2009-2011	式(4)	40%	(0.72) 33%	(1.00) 35%
		(1.00)	(0.72) (0.21)	(0.36) (0.51)
2012	式(5)	40%	(0.72) 52%	(0.64) 44%

※各セルの値の左、上、左上に記述した()内に、左、上、左上の各セルの値との差の検定結果 (p 値) を示す。p ≤ 0.10 の場合、危険率10%で有意差があると判断し、p 値を下線で強調する。

4.2 データ追加収集による精度変化

データ追加収集による精度変化を表7、表8の列方向に示す。モデルパラメータを固定し、データを変化させた場合の精度変化を、左から右へ古い年度から順に記す。各年度の値について、MdMREではWilcoxonの順位検定[27]、Pred25ではFisherの正確確率検定[27]で前年度の値との差を検定した。各セルの値の左側に記す()内に、直左のセルに記す前年度の値との差の検定結果 (p 値) を示す。p ≤ 0.10 の場合、危険率10%で有意差があると判断し、p 値を下線で強調する。

たとえば、表7の最上行中央列のセルは、2009～2011年度のデータ33件を分析し、2008年度のモデルパラメータで式(3)の重回帰モデルを作成した結果である。0.57は、33件のプロジェクトについて算出したMdMREである。当該セルの値の左に記す(0.08)は、2008年度のデータ10件を分析した場合のMdMRE 0.31との差の検定結果 (p 値) である。ここでは、p ≤ 0.10 であり、危険率10%で有意差があると判断できるため、0.08を下線で強調している。

検証の結果、モデルパラメータを2008年度に固定した場合、2009～2011年度の大規模なデータ増加でMdMRE、Pred25ともに精度が大きく悪化した (MdMREは統計的有意差も観察できた)。この理由として、2009～2011年度の追加収集データに、2008年度より多様なプロジェクトが含まれたことが考えられる。表5に示すように、2009～2011年度のデータは2008年度に比べて基盤構築工数と論理サーバ台数の最大値と変動係数が大きく増加している。これにより、3章(6)で説明したように適用可能プロジェクトを拡大できた一方、モデルパラメータを見直さなければ、精度が大きく悪化してしまったと考えられる。

2012年度のデータ追加収集は、2009～2011年度ほど大きな精度の変化はなく、MdMRE、Pred25ともに統計的有意差は観察できなかった。表5に示すように、基盤構築工数、論理サーバ台数、基盤構築期間いずれも最大値や変動係数に前年度からの大きな変化はない。2012年度は、データが10件程度増加したが、データに含まれるプロジェクトの多様性は大きく変化しなかったと考えられる。

4.3 パラメータ見直しによる精度変化

パラメータ見直しによる精度変化を表7、表8の行方向に示す。学習データおよびテストデータを固定し、モデルパラメータを変化させた場合の精度変化を、上から下へ古い年度から順に記述する。各年度の値について、MdMREではWilcoxonの符号順位検定[27]、Pred25ではFisherの正確確率検定[27]で前年度の値との差を検定した。各セルの値の上に記す()内に、直上のセルに記す前年度の値との差の検定結果 (p 値) を示す。p ≤ 0.10 の場合、危険率10%で有意差があると判断し、p 値を下線で強調する。

検証の結果、データを固定した場合、すべての年度の

データにおいて、多くの場合に MdMRE, Pred25 ともに精度が改善した。特に、2009~2011 年度から 2012 年度のパラメータに変更した場合には MdMRE で統計的有意差も観察できた。2008 年度から 2009~2011 年度のパラメータ変更 (式 (3) から式 (4)) では、2 値のカテゴリ変数であるアーキテクチャ経験有無から、3 値の非機能要求グレードに変更している。さらに、2012 年度のパラメータ変更 (式 (4) から式 (5)) では比率尺度の開発期間を追加している。モデルが含む情報量の増加に従って精度が改善し、また 3 章 (8) で説明した多重共線性の発生もない。パラメータを適宜妥当に見直すことで、継続的に精度を改善できたと考える。

4.4 データ追加収集とパラメータ見直しによる精度変化

データ追加収集とパラメータ見直しによる精度変化を表 7, 表 8 の斜め方向に示す。学習データおよびテストデータとモデルパラメータの両方を変化させた場合の精度変化を、左上から右下へ古い年度から順に記述する。各年度の値について、MdMRE では Wilcoxon の順位和検定 [27], Pred25 では Fisher の正確確率検定 [27] で前年度の値との差を検定した。各セルの値の左上に記す () 内に、直左上のセルに記す前年度の値との差の検定結果 (p 値) を示す。p ≤ 0.10 の場合、危険率 10% で有意差があると判断し、p 値を下線で強調する。

検証の結果、統計的に有意差は観察できなかったが、2009~2011 年度は MdMRE, Pred25 ともに精度が悪化し、2012 年度は改善した。この理由として、表 5 に示すように、2009~2011 年度のデータは 2008 年度に比べて最大値や変動係数が大きく増加したが、モデルパラメータは式 (3) から式 (4) のように実質的にはダミー変数を 1 つ追加しただけであったことがあげられる。データに含まれるプロジェクトの多様性増大に対して、重回帰モデルの見積り精度改善が小さかったため、2009~2011 年度は精度が低下したと考えられる。一方、2012 年度のデータは表 5 に示すように、見積り対象の工数やパラメータの最大値や変動係数の増加はほぼなかったのに対し、式 (5) のように比率尺度のパラメータを追加したため、精度が改善したと考えられる。

5. 考察

5.1 既存研究との見積り精度比較による実用性

見積りモデルおよびツールの実用性を考察するため、既存研究で報告された精度との比較結果を表 9 に示す。表中、最上行から下へ順に、当取り組みにおける各年度のモデル、2.2 節で紹介した金子ら [21] と Valerdi [37] のモデル、2.1 節で紹介した Moløkken ら [26] の調査で対象とされた文献 [2], [18], [30], [35] で報告された精度をデータ件数とともに示す。なお、2.2 節で紹介した小澤 [29] の研究

表 9 既存研究との見積り精度の比較

Table 9 Comparison of estimation accuracy with past researches.

	見積対象	データ件数	MRE	Pred25
2008		10	(中央値) 0.31	40%
2009-2011	基盤構築	33	(中央値) 0.46	33%
2012	工数	43	(中央値) 0.34	44%
金子ら [21]		5	(中央値) 0.02	60%
Valerdi [37]	システム 開発工数	27	—	27%
Jenkins ら[18]	アプリケ ーション	72	(中央値) 0.34	—
Phan [30]		191	(平均値) 0.33	—
Bergeron ら[2]	開発工数	89	(平均値) 0.33	—
Standish [35]		365	(平均値) 0.89	約 20%

※—は、当該指標や実験結果データの報告がないことを示す。

や Moløkken らの調査における他の研究では、精度が報告されていないため比較できなかった。

比較の結果、2012 年度に見積りモデルを実用に耐えうる精度に改善できたと考えられる。表 9 に示す文献の多くでは、MRE の中央値または平均値は 0.33~0.34 であり、Pred25 は 20~27% であった。当取り組みのモデルは 2012 年度にある程度のデータ件数によって統計的妥当性を担保したうえで、これと同程度の精度を達成できた。金子らの研究と比べると精度が低く見えるが、金子らの研究ではデータ件数が非常に少なく、過適合の発生など統計モデルとしての妥当性には疑問がある。一方、当取り組みでは、2012 年度には、統計モデルとしての妥当性を十分担保できるデータ件数を確保できている。

5.2 見積り誤りが大きかったプロジェクトとその原因

2012 年度には全体として実用に耐えうる精度を達成できた一方、他と比べて見積り誤りが大きかったプロジェクトも少数ではあるが存在した。同一構成の基盤を、多数のシステム稼働拠点に複製構築する場合に、本稿の見積りモデルはやや過大見積りとなった。基盤の設計や構築手順を再利用できるため、論理サーバ台数 1 台あたりの設計、構築、テスト工数が他より小さかったためであると考えられる。また、システムがデータをやり取りする外部接続システムが多い場合に、本稿の見積りモデルはやや過小見積りとなった。論理サーバ台数 1 台あたりの工数、並びに、各外部接続システムとの連携確認テストの工数が大きかったためであると考えられる。

5.3 利用者数の増加と利用者からのフィードバック

表 6 に示すように、2012 年度までは毎年前年度の約 2 倍程度まで利用者は順調に増加し、3 章 (10) で示したとおり、取り組みに対して肯定的な意見が大半であった。特

に、基盤構築に関わる実務者からは、取り組みへの肯定的な意見が多かった。分析データの収集についても、多くのプロジェクトが協力的であり、取り組みの継続実施を求める意見が大半であった。理由として、アプリケーション開発の見積りについては多くの研究が実施されている [1], [3], [4], [7], [17], [31], [33], [39] のに対し、基盤構築の見積りについては研究が少なく、実務者が自身の見積りの妥当性を検証することが難しい実態があると考えられる。2.2 節でも紹介したとおり、業務の一環として基盤構築を行う企業で研究が行われているものの [21], [22], [29], アプリケーション開発の見積りと比較すると、取り組みの規模や数が十分であるとはいいいにくい。ニーズを抱える企業側でも取り組みを継続するとともに、大学や研究機関での研究推進も必要であると考えられる。

6. おわりに

本稿では、見積りモデルを企業のシステム開発実務で利用している事例として、NTT データにおける基盤構築工数見積りモデルの継続的改善と普及展開について報告した。2008～2012 年度の取り組みを紹介し、取り組みによる精度の改善について実プロジェクトのデータを利用して検証した。また、モデルの精度を既存研究と比較して実用性を考察し、実務者から得たフィードバックを紹介した。

2014 年度現在、本稿で報告した基盤構築工数見積りモデルに加えてアプリケーション開発工数見積りモデルを作成し、両者を 1 つの Web アプリケーションに組み込んで国内の全グループ会社にサービス提供している。全グループ会社の社員は誰でも当該サービスを利用可能であり、大幅な見積り誤りがないかの検証や、生産性向上効果の検証に利用されている。また、見積り誤りに起因する問題化プロジェクト発生防止のための第三者見積り審査への応用も検討されている。

今後の課題として、SAP などのパッケージ製品を利用した開発、Hadoop などを利用する大規模クラスタ基盤、仮想化技術を利用するクラウド基盤など、技術的に多様化する基盤構築方法に対応するため、それらの構築プロジェクトからデータを収集し、モデルを改善すること。また、国内だけでなく海外グループ会社からもデータを収集し、適用可能プロジェクトの拡大に応じてモデルパラメータを継続的に見直すとともに、開発した Web アプリケーションのサービスを長期的な戦略の下に改善していくことがあげられる。

謝辞 本研究の実施にあたり、実績データ収集のためのアンケートにご協力いただいたプロジェクトの皆様へ深く感謝する。

参考文献

[1] Albrecht, A.J. and Gaffney, J.E.: Software Function, Source Lines of Code, and Development Effort Pre-

- diction, *IEEE Trans. Softw. Eng.*, Vol.9, pp.639-648 (1983).
- [2] Bergeron, F. and St-Arnaud, J.-Y.: Estimation of Information Systems Development Efforts: A Pilot Study, *Information & Management*, Vol.22, pp.239-254 (1992).
- [3] Boehm, B.W., Madachy, R. and Steece, B.: *Software Cost Estimation with Cocomo II*, Prentice Hall PTR (2000).
- [4] Briand, L.C., Basili, V.R. and Thomas, W.M.: A Pattern Recognition Approach for Software Engineering Data Analysis, *IEEE Trans. Softw. Eng.*, Vol.18, pp.931-942 (1992).
- [5] Cohen, J.: *Statistical Power analysis for the Behavioral Sciences*, 2nd ed., Lawrence Erlbaum, Hillsdale, NJ (1988).
- [6] Connolly, T. and Dean, D.: Decomposed versus Holistic Estimates of Effort Required for Software Writing Tasks, *Management Science*, Vol.43, No.7, pp.1029-1045 (1997).
- [7] Delaney, W.A.: Predicting the Costs of Computer Programs, *Data Processing Magazine*, pp.32-34 (1966).
- [8] Dobson, A.J. (著), 田中 豊, 森川敏彦, 山中竹春, 富田誠 (訳): 一般化線形モデル入門, pp.115-116, 共立出版 (2008).
- [9] (独) 情報処理推進機構ソフトウェア・エンジニアリング・センター (IPA-SEC): ソフトウェア開発見積りガイドブック—IT ユーザとベンダにおける定量的見積りの実現 (2006).
- [10] (独) 情報処理推進機構ソフトウェア・エンジニアリング・センター (IPA-SEC): ソフトウェア改良開発見積りガイドブック—既存システムがある場合の開発 (2007).
- [11] (独) 情報処理推進機構ソフトウェア・エンジニアリング・センター (IPA-SEC): ソフトウェア改良開発見積りガイドブック—品質要件に応じた見積りとは (2008).
- [12] (独) 情報処理推進機構ソフトウェア・エンジニアリング・センター (IPA-SEC): 非機能要求の見える化と確認の手段を実現する「非機能要求グレード」の公開 (2010).
- [13] Heemstra, F.J. and Kusters, R.J.: Controlling Software Development Costs: A Field Study, *International Conference on Organisation and Information Systems*, Bled, Yugoslavia (1989).
- [14] Hill, J., Thomas, L.C. and Allen, D.E.: Experts' Estimates of Task Durations in Software Development Projects, *International Journal of Project Management*, Vol.18, No.1, pp.13-21 (2000).
- [15] Hoch, S.J. and Schkade, D.A.: A Psychological Approach to Decision Support Systems, *Management Science*, Vol.42, No.1, pp.51-64 (1996).
- [16] 井ノ口伸人, 伏田享平, 渡辺絢子, 大杉直樹, 藤貫美佐, 渡辺真太郎, 戸村元久, 木谷 強: システム基盤構築の工数見積りモデル・ツール開発と普及展開, SQIP2012 (2012).
- [17] Jørgensen, M.: Experience with the Accuracy of Software Maintenance Task Effort Prediction Models, *IEEE Trans. Softw. Eng.*, Vol.21, No.8, pp.674-681 (1995).
- [18] Jenkins, A.M., Naumann, J.D. and Wetherbe, J.C.: Empirical Investigation of Systems Development Practices and Results, *Information & Management*, Vol.7, pp.73-82 (1984).
- [19] Jørgensen, M.: A Review of Studies on Expert Estimation of Software Development Effort, *Journal of Systems and Software*, Vol.70, No.1-2, pp.37-60 (2004).
- [20] Kahneman, D., Slovic, P. and Tversky, A.: *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge University Press, Cambridge, United Kingdom (1982).

- [21] 金子美和, 伊藤 稯, 田中正美, 弘末太郎: 大規模 IT インフラ構築プロジェクトにおける概算工数見積もりモデル, PM 学会研究発表大会予稿集 2008 (春季), pp.111-114 (2008).
- [22] Kaneko, M.: Model in Approximate Estimating Effort for Large IT Infrastructure Building Projects, *Proc. Pro-MAC 2010*, pp.1301-1308 (2010).
- [23] Lederer, A.L. and Prasad, J.: Nine Management Guidelines for Better Cost Estimating, *Comm. ACM*, Vol.35, No.2, pp.51-59 (1992).
- [24] Lim, J.S. and O'Connor, M.: Judgmental Forecasting with Time Series and Causal Information, *International Journal of Forecasting*, Vol.12, No.1, pp.139-153 (1996).
- [25] McAulay, K.: Information Systems Development and the Changing Role of MIS in the Organisation, *1st New Zealand MIS Management Conference*, Wellington (1987).
- [26] Moløkken, K. and Jørgensen, M.: A Review of Surveys on Software Effort Estimation, *Proc. 2003 Intl. Symposium on Empirical Soft. Eng. (ISESE 2003)*, pp.223-230 (2003).
- [27] Myles, P.S. and Gin, T.: *Statistical Methods for Anaesthesia and Intensive Care*, Butterworth-Heinemann (2000).
- [28] O'Connor, M., Remus, W. and Griggs, K.: Judgmental Forecasting in Times of Change, *International Journal of Forecasting*, Vol.9, No.2, pp.163-172 (1993).
- [29] 小澤良男: システム基盤設計構築における見積もり手法の確立, NRI 技術創発, Vol.9, pp.70-85 (2007).
- [30] Phan, D.: Information Systems Project Management: An Integrated Resource Planning Perspective Model, *Department of Management and Information Systems*, Tucson, Arizona (1990).
- [31] Putnam, L.H.: A General Empirical Solution to the Macro Software Sizing and Estimation Problem, *IEEE Trans. Softw. Eng.*, Vol.4, pp.345-381 (1978).
- [32] Shanteau, J.: Competence in Experts: The Role of Task Characteristics, *Organizational Behaviour and Human Decision Processes*, Vol.53, No.2, pp.252-266 (1992).
- [33] Shepperd, M. and Shfield, C.: Estimating Software Project Effort Using Analogies, *IEEE Trans. Softw. Eng.*, Vol.21, pp.126-137 (1997).
- [34] Soll, J.B.: Determinants of Overconfidence and Miscalibration: The Roles of Random Error and Ecological Structure, *Organizational Behaviour and Human Decision Processes*, Vol.65, No.2, pp.117-137 (1996).
- [35] Standish Group: The Chaos Report, The Standish Group (1994).
- [36] (社) 日本情報システム・ユーザ協会: ソフトウェア開発管理基準に関する調査報告書 (ソフトウェアメトリックス調査) (2012).
- [37] Valerdi, R.: *The Constructive Systems Engineering Cost Model (COSYSMO): Quantifying the Costs of Systems Engineering Effort in Complex Systems*, VDM Verlag (2008).
- [38] Whitecotton, S.M., Sanders, D.E. and Norris, K.B.: Improving Predictive Accuracy with A Combination of Human Intuition and Mechanical Decision Aids, *Organizational Behaviour and Human Decision Processes*, Vol.76, No.3, pp.325-348 (1998).
- [39] Wolverton, R.W.: The Cost of Developing Large-Scale Software, *IEEE Trans. Softw. Eng.*, Vol.C-23, No.6, pp.615-636 (1974).
- [40] Wydenbach, G. and Paynter, J.: Software Project Estimation: A Survey of Practices in New Zealand,

New Zealand Journal of Computing, Vol.6, pp.317-327 (1995).



井ノ口 伸人

ストの研究に従事。

株式会社 NTT データ。2005 年名古屋大学工学部電気電子・情報工学科卒業。2007 年同大学大学院情報科学研究科博士前期課程修了。同年株式会社 NTT データ入社。現在, エンピリカルソフトウェア工学, ソフトウェアテ



大杉 直樹 (正会員)

エンピリカルソフトウェア工学, 生産性/品質データ収集/分析, メトリックスの研究に従事。電子情報通信学会, IEEE 各会員。

株式会社 NTT データ。2001 年奈良工業高等専門学校専攻科・電子情報修了。2004 年奈良先端大情報科学研究科博士後期課程修了。同年同大学特任助手。2007 年株式会社 NTT データ入社。技術開発本部勤務。博士 (工学)。



伏田 享平 (正会員)

データ入社。現在, 技術開発本部プロジェクトマネジメント・イノベーションセンタ所属。博士 (工学)。ソフトウェア工学, 特にソフトウェア開発データの分析・利用に関する研究に従事。IEEE, ソフトウェア科学会各会員。

株式会社 NTT データ。2005 年大阪府立大学工学部電気電子システム工学科中退。2010 年奈良先端科学技術大学院大学情報科学研究科博士後期課程修了。同年同大学博士研究員。2011 年同特任助教。2012 年株式会社 NTT



渡辺 絢子

株式会社 NTT データ。2008 年早稲田大学理工学部社会環境工学科卒業。2010 年東京大学大学院工学系研究科社会基盤学専攻修了。同年株式会社 NTT データ入社。統計的見積りの研究に従事。



吉野 順 (正会員)

株式会社 NTT データ。1989 年東京電機大学大学院理工学研究科修士課程修了。同年 NTT データ通信株式会社 (現, 株式会社 NTT データ) 入社。技術開発本部勤務。2005 年よりソフトウェア開発業務, 事務処理オペレー

ション業務における作業の見える化, 生産性および品質改善に向けての研究開発に従事。



木谷 強 (正会員)

株式会社 NTT データ。1983 年慶應義塾大学工学部電気工学科卒業。同年日本電信電話公社入社, 1988 年株式会社 NTT データへ転籍。自然言語処理, 情報検索, ソフトウェア工学の研究開発に従事。博士 (工学)。



藤貫 美佐

株式会社 NTT データ。1987 年東京女子大学文理学部数学科卒業。同年日本電信電話株式会社入社。ミドルウェアの開発, 社内開発標準の策定, FP 法の普及展開, 統計的見積りの研究に従事。日本ファンクションポイントユー

ザ会副会長兼事務局長。



渡辺 真太郎 (正会員)

株式会社 NTT データ。1994 年東京工業大学理学部情報科学科卒業。1996 年奈良先端科学技術大学院大学博士前期課程了。同年 NTT データ通信株式会社 (現, 株式会社 NTT データ) 入社。現在に至る。M2M/IoT ソリュー

ションの研究開発に従事。博士 (工学)。電子情報通信学会会員。



戸村 元久

株式会社 NTT データ。1987 年慶應義塾大学理工学部電気工学科卒業。1989 年同大学大学院理工学研究科修士課程修了。同年株式会社 NTT データ入社。情報システム生産技術の開発に従事。