

THE IEICE TRANSACTIONS ON INFORMATION AND SYSTEMS (JAPANESE EDITION)

IEICE | **電子情報通信学会**
D | **論文誌** 情報・システム

VOL. J103-D NO. 3

MARCH 2020

本PDFの扱いは、電子情報通信学会著作権規定に従うこと。

なお、本PDFは研究教育目的（非営利）に限り、著者が第三者に直接配布することができる。著者以外からの配布は禁じられている。

情報・システムソサイエティ

一般社団法人 **電子情報通信学会**

THE INFORMATION AND SYSTEMS SOCIETY

THE INSTITUTE OF ELECTRONICS, INFORMATION AND COMMUNICATION ENGINEERS

系統的文献レビューを用いたファンクションポイント研究の調査*

山田 涼太[†] 楠本 真二[†] 肥後 芳樹[†] 杉本 真佑[†]
倉重 誠^{††}

Survey of Function Point Studies Using Systematic Literature Review*

Ryota YAMADA[†], Shinji KUSUMOTO[†], Yoshiki HIGO[†], Shinsuke MATSUMOTO[†],
and Makoto KURASHIGE^{††}

あらまし 近年、ソフトウェアの規模を評価する尺度としてファンクションポイント (FP) の重要性が高まってきた。FP に関連してどのような研究がこれまでにされているかを明らかにするため、本論文では系統的文献レビュー (Systematic Literature Review) を実施した。ソフトウェア工学分野における主要な論文誌や国際会議において、1979 年から 2018 年までに発表された文献のうち 95 本を調査した。まず、FP に関連したテーマを 6 種類に分類し、それぞれのテーマについて研究動向をまとめた。その結果、初期の 20 年では、他の開発規模尺度との比較や工数見積り等への活用、直近 20 年では FP 自動計測のテーマが主流であることがわかった。次に、FP は開発現場での実用性が重視されることから、研究成果の開発現場での適用可能性や開発現場のニーズと研究とのギャップについて調査し、今後の課題についてまとめた。

キーワード ファンクションポイント, 系統的文献レビュー, ソフトウェア工学, 見積り, IFPUG 法

1. ま え が き

ソフトウェア開発の見積りの元となる開発規模の尺度として、ファンクションポイント (FP) が存在する [2]。FP はソフトウェアの各機能の複雑さを元にソフトウェアの開発規模を測定する手法であり、ユーザ視点で機能を計測する、実装依存しない等の利点が指摘されている。我が国においても、政府情報システムの開発で、原則 FP の見積りとその根拠を示すことが必須となっている [24]。

一方で、FP には、主観的判断が含まれる、計測の手間が大きい等の問題も指摘されており、この 40 年間で様々な研究が行われてきている。研究対象の範囲は多岐に渡り、その論文数は膨大となっているため、新規の研究者や開発者が FP に既存研究を把握する事が難しい状況である。また、FP は見積り等、開発現

場における実用性が重視される技術であるため、研究成果の現場への適用可能性を知る事は重要である。更に、近年、ソフトウェア工学研究分野において開発現場のニーズと研究の間にギャップがあるという指摘がされている [11]。FP に関しても研究と現場のニーズの関係を確認する事は非常に重要である。

以上の背景から、本研究では過去 40 年間に発表された研究論文を対象に系統的文献レビュー (Systematic Literature Review) を実施した。系統的文献レビューとは、網羅的で再現性のある文献調査手法である [5], [17], [19]。系統的文献レビューは主に、新規研究者の補助を目的とした既存研究の要約や既存の研究成果から知見を得るために実施される。

本論文では、以下の三つの Research Question を設定し、ソフトウェア工学分野における主要な論文誌や国際会議において採録された論文 95 本を対象に系統的レビューを実施した。

(RQ1) FP 研究の動向はどのようなものか

(RQ2) 研究成果は現場での適用が可能か

(RQ3) 開発現場のニーズと研究にギャップは存在するか

本レビューによってこれまでの FP 研究を系統的に整理し、開発現場における実用性や現場と研究のギャ

[†] 大阪大学大学院情報科学研究科, 吹田市
Graduate School of Information Science and Technology,
Osaka University, 1-5 Yamada-oka, Suita-shi, 565-0781
Japan

^{††} (株) SOFTEST, 川崎市
SOFTEST, Co., Ltd., Kawasaki-shi, 211-0005 Japan

* 本論文は、システム開発論文である。

DOI:10.14923/transinfj.2019JDR0001

プを明らかにすることを旨とする。その結果は、研究者や実務者にとって、関連研究調査の参考になるとともに、研究者の今後の研究の方針を考える上での参考になる。

以降、**2.**では本研究の前提知識について簡単に説明する。**3.**では系統的文献レビューの手順について述べ、**4.**では系統的文献レビューを実施して得られた情報から本三つのRQへの回答を行う。**5.**ではFP研究における今後の課題の考察を行い、**6.**でまとめる。

2. 準備

ここでは本研究の用語や関連研究を簡単にまとめる。

2.1 ファンクションポイント法

ファンクションポイント (FP) 法 [2] はソフトウェアの規模を計測する手法の一つである。FP 法は、ユーザから見た機能量を計測する手法で、Albrecht によって 1979 年に提案された [2]。画面や帳票、ファイルなどを通じた情報の入出力に着目し、それらを種類別に数え上げ、種類数を加重合計した値を機能量としている。これまでに、IFPUG 法 [13]、MarkII 法 [25]、COSMIC 法 [26] 等、様々な FP 法が提案されてきている。

以下では、特に広く普及している IFPUG 法を紹介する。IFPUG 法は、Albrecht 版の使い勝手や曖昧な部分を改良したバージョンであり、ビジネスアプリケーションソフトウェアを対象として欧米で広く使用されている。

IFPUG 法では、ソフトウェアのもつ機能を次の 5 種類の基本機能要素に分類する [13]。

- 内部論理ファイル (ILF)

計測対象のアプリケーション内でデータが更新される論理的な関連をもったデータの集合

- 外部インタフェースファイル (EIF)

計測対象のアプリケーションによってデータが参照されるデータの集合 (データは更新されない)

- 外部入力 (EI)

計測境界外からのデータ入力によって ILF の更新を行う処理

- 外部出力 (EO)

計測境界外へのデータ出力を含む処理のうち、出力時に何らかの加工を伴うもの

- 外部照会 (EQ)

計測境界外へのデータ出力を含む処理のうち、ILF や EIF の内容をそのまま出力するものであり、処理が

ILF を更新しないもの

それぞれの機能に対し、扱うデータ項目の数などを元に重みづけを行い、その合計をソフトウェアの開発規模 (FP) として扱う。開発環境や開発言語等に依存せず、ユーザから見た機能数を測るため、ユーザ側から見てもわかりやすいという利点があり、契約時のベンダとの合意やプロジェクト間の生産性評価等での利用で有用であると言われている。

2.2 Systematic Literature Review

系統的文献レビュー (Systematic Literature Review, SLR) とは、網羅的で再現性のある文献調査手法である [5], [17], [19]。SLR では明確な手順を踏んで調査を行うため、レビューの再現性が高い。また、手順の設定が明確に行われていれば同様の結果を得る事ができ、レビュー結果の信頼性も高いと言われている。特に医療の研究分野では、情報収集や既存症例の調査を行う際によく用いられる手法である。ソフトウェア工学分野では、SLR のガイドラインが Kitchenham らにより提案されており [17]、研究で設定した仮説の検証や既存研究の要約を行うために実施されることが多い [18]。本論文で実施する SLR も Kitchenham らが提案した手法に従う。

SLR の具体的な手順については **3.** で述べる。

2.3 Zelkowitz の評価モデル

Zelkowitz の評価モデルとは、コンピュータサイエンス分野における研究の評価方法を 12 種類に分類したものである [27]。以下に示すように、大分類として次の (1)~(3) の 3 種類、小分類として 12 種類の分類が存在する。

(1) Observational Methods: 研究成果を企業や大学などのプロジェクトに適用する事でその有用性を評価する。

- Project monitoring: 単一のプロジェクトに研究成果を適用し、その結果を評価する。

- Case study: 単一のプロジェクトに研究成果を適用し、従来手法などの結果と比較する事でその有用性を評価する。

- Assertion: ある規模以下の複数のプロジェクトに研究成果を適用し、従来手法での結果と比較する事で、研究成果が有効であるかどうかを評価する。

- Field study: 複数のプロジェクトに研究成果を適用し、従来手法などの結果と比較する事で、研究成果が一般的に有効であるかを評価する。

(2) Historical Methods: 既に完了しているプロジェ

クト等の既存データから知見を得る。

- Literature search: 過去の文献を元に知見を得る。SLR が当てはまる。

- Study of Legacy: 大量の過去プロジェクトを分析する事で知見を得る (主に、工数や開発規模のような数値的なデータを分析する)。

- Study of Lessons-learned: 過去プロジェクトを実行した際に得られた情報や開発者へのインタビュー等から、プロジェクトや技術に対する教訓を収集し、まとめる。

- Static analysis: 完成したプロジェクトの成果物を分析する事で知見を得る (主に、仕様書やソースコードのような成果物を分析する)。

(3) Controlled Methods: 研究成果の統計的妥当性を示すために、Observational Methods 以上に厳密な比較・評価を行う。

- Replicated Experiment: 複数のプロジェクトに研究成果を含めた複数の方法を適用し、その結果を比較する事で研究成果の有用性などを評価する。

- Synthetic Environment Experiments: 複数の被験者の半分が提案手法を、もう半分が従来手法を用い、その結果を比較するなどして研究成果の有用性などを評価する。

- Dynamic analysis: プログラムの性能などを比較する事で、研究成果の有用性などを評価する。

- Simulation: シミュレーションによって研究成果の有用性などを評価する。

本研究ではこの評価モデルを用いて、FP の研究成果の開発現場での適用可能性を評価する。

3. 実施した系統的文献レビュー

実施した系統的文献レビューの各ステップについて、以降で説明する。

3.1 Research question の設定

本論文では以下の三つの Research question を設定した。

(RQ1) FP 研究の動向はどのようなものか: 既存研究の研究動向を明らかにするため、既存研究にはどのようなトピックが存在するかを調査し、研究内容がどのように推移したかを分析する。

(RQ2) 研究成果は現場での適用が可能か: 研究成果の開発現場での適用可能性を確認するため、既存研究の研究成果がどのように評価されているかを調査する。

(RQ3) 開発現場のニーズと研究にギャップは存在す

るか: FP に関する開発現場のニーズと研究の間のギャップを調査し、今後の研究の方向性を探る。

3.2 研究論文の収集

調査対象となる研究論文の収集を行い、調査対象論文の候補とする。収集の際には、あらかじめ検索期間と検索対象データベース、及び検索に用いるキーワードを設定する。

検索する期間は FP が提案されてから今日まで (1979 年から 2018 年) とした。検索対象データベースは、ソフトウェア工学分野において著名な論文誌や国際会議の論文集に採録されている研究論文が網羅されている点から、IEEE xplore [12], ACM Digital library [1], Science Direct [22], Scopus [23], Google Scholar [9] の五つを選択した。検索に用いるキーワードとして “function point”, “function points”, “function-point”, “function-points” を適用した。

3.3 研究論文の選別

各検索対象データベースから収集した論文候補の選別を行い、調査対象となる研究論文を決定する。選別の手順を以下に示す。

(1) 各論文のタイトルを確認し、重複している論文を候補から取り除く。

(2) 各論文のタイトル、内容梗概、キーワードを確認し、ソフトウェア工学分野に属しない論文を候補から取り除く。

(3) 各論文の出典を確認し、CORE RankingsPortal [6] においてランクが A 以上である論文誌または国際会議の論文集及び H index [10] が 80 以上である論文誌または国際会議の論文集に採録されている論文以外を候補から取り除く。

(4) 残った候補を調査対象論文として選択する。

上記選別手順を図 1 に示す。選別の結果、95 本が調査対象の論文として選択された。調査対象の論文は参考文献の [28]~[122] に対応する。掲載元別の集計表を表 1 に示す。

3.4 情報の抽出

調査対象とした研究論文の本文を確認し RQ の回答に必要な情報として主に以下の情報を抽出した。

(1) タイトル, (2) 著者名, (3) 出典名, (4) 出版年, (5) 研究の背景, (6) 研究の目的, (7) 手法, (8) 結果, (9) 結論。

主に, (4)(5)(6)(7)(9) は RQ1, (8) は RQ2, (9) は RQ3 の回答のためにそれぞれ抽出した。

表 1 掲載元別の集計表
Table 1 Source and number of papers.

| Journal/Conference | Number |
|--|--------|
| Information and Software Technology | 20 |
| Journal of Systems and Software | 14 |
| International Symposium on Empirical Software Engineering and Measurement | 10 |
| IEEE Transactions on Software Engineering | 10 |
| International Conference on Software Engineering | 5 |
| International Symposium on Software Metrics | 5 |
| Expert Systems with Applications | 3 |
| Communications of the ACM | 2 |
| IEEE Computer | 2 |
| Euromicro Conference Software Engineering and Advanced Applications | 2 |
| Industrial Management and Data Systems | 2 |
| IEEE Software | 2 |
| Americas Conference on Information Systems | 1 |
| The Computer Journal | 1 |
| Evaluation and Assessment in Software Engineering | 1 |
| European Journal of Information Systems | 1 |
| International Conference on Advanced Information Systems Engineering | 1 |
| International Conference on Data Mining Workshops | 1 |
| International Conference on Engineering of Complex Computer Systems | 1 |
| International Conference on Information Systems | 1 |
| International Conference on Model Driven Engineering Languages and Systems | 1 |
| International Conference on Software and Systems Process | 1 |
| Information Sciences | 1 |
| Information Systems Journal | 1 |
| Journal of Computer Information Systems | 1 |
| Journal of Information Technology | 1 |
| Journal of Management Information Systems | 1 |
| Journal of Neuroscience Methods | 1 |
| Working Conference on Mining Software Repositories | 1 |
| IEEE Transactions on Engineering Management | 1 |
| Total | 95 |

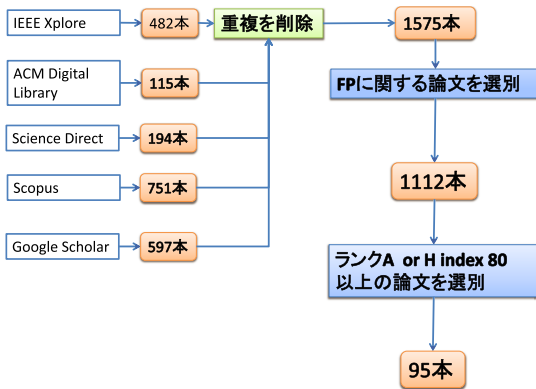


図 1 論文収集と選別の手順

Fig. 1 Procedure of collecting and sorting papers.

3.5 RQ への回答方法

抽出した情報を整理・考察した上で (RQ1)~(RQ3) に対して、下記の方針で回答を行う。

(RQ1) に対しては、95 本の論文をその内容から幾つ

かのトピックに分類した上で、過去 40 年間の研究を 5 年ごとに分け、年代頃の本数と各トピックの割合から研究内容の推移を分析する。(RQ2) に対しては、各研究で用いられている評価を 2.3 で紹介した Zelkowitz の評価モデルを用いて分類し、結果の現場での適用可能性について分析する。(RQ3) に対しては、開発現場のニーズと実際に行われている研究にギャップが存在するかを確認するために、JFPUG (日本ファンクションポイントユーザ会) の FP 活用研究会で得られた FP に対する開発現場のニーズを利用した。JFPUG は IFPUG (国際ファンクションポイントユーザ会) の日本支部であり、日本国内における FP の理解と普及、その利用法の検討及びソフトウェア計測の検討を主な目的としている。また、FP 活用研究会は現場における FP 活用の課題に対する解決策を探求し、定量データを利用する様々な活動の改善につなげることを目的としている。得られた現場のニーズと既存研究を照らし合わせることで、ギャップの分析を行う。

各集計結果と (RQ1)~(RQ3) への回答については 4. で、結果の考察については 5. で、それぞれ説明する。

4. レビュー結果

4.1 RQ1

4.1.1 研究トピック

95 本の論文を六つのトピック (T1)~(T6) に分類した。分類された論文とその本数を表 2 に示す。ここでは、95 本の各論文をそれぞれいずれか一つのトピックに分類している。論文によっては二つのトピックに関係した論文もあったが、その場合は内容を吟味し、二つのトピックの中で主たる方のトピックを分類先として選んだ^(注1)。

(T1) FP の利点や欠点の評価: FP の計測精度や信頼性を評価する研究。研究成果の評価ではなく、既に存在している FP 法についての評価が行われている研究がこのトピックに含まれる。

(T2) 計測補助: FP 計測の手間を削減するための自動計測ツールなどの研究。計測者が手動で行っていた FP 計測の補助を行う研究がこのトピックに含まれる。FP 計測が難しいとされている開発での計測について

は後述の (T5) に含まれる。

(T3) 計測結果の活用: 主に計測された FP を用いて工数や生産性を導出するための研究。FP の結果を見積りなどに活用するための研究がこのトピックに含まれる。

(T4) 計測ルールの変更: FP の計測ルールの改善若しくは新たな計測ルールの提案についての研究。計測要素や重みづけの再定義など、既存の計測ルールの改善案を提案している研究がこのトピックに含まれる。

(T5) 計測が難しいとされている対象への適用: FP の適用が難しいとされる開発手法やアプリケーションへの適用についての研究。FP の適用が難しいとされている非ウォーターフォール型開発に対する計測方法の提案などがこのトピックに含まれる。既に計測方法が確立されているウォーターフォール型開発への適用などは (T1) に含まれ、計測の手間の削減などの計測補助が主目的である研究は (T2) に含まれる。

(T6) その他: 異なる FP 法で計測された FP の値の変換や FP 研究についての SLR 等 (T1)~(T5) に当てはまらない研究がこのトピックに含まれる。

4.1.2 研究内容の推移

次に、FP 研究における研究内容の推移について分析を行った。分析のために、40 年間の研究を 5 年ごとに分け、年代ごとの本数と各トピックの割合から研究内容の推移を分析した。トピックごとに色分けした出版年代ごとの調査対象論文数を図 2 に示す。

初期を除いて、六つのトピックの研究は継続的に行われていることが分かる。

前半 20 年、特に 1989 年以降において (T1) と (T3) についての研究が多くなっている。これは FP が提案された 1979 年以降に FP が普及し、その利点や活用方法に関心が集まったためと考えられる。

後半 20 年においては 1999 年から 2003 年において

表 2 各トピックに含まれる論文と本数

Table 2 Articles and number of articles included in each topic.

| トピック | 含まれる論文 | 本数 |
|-------------------------|---|----|
| (T1) FP の利点や欠点の評価 | [28], [30], [32], [33], [36], [37], [39]~[42], [44], [45], [47], [48], [51], [54], [56], [63], [66], [77], [81], [90], [92], [95], [101], [105], [108], [118] | 28 |
| (T2) 計測補助 | [55], [57], [65], [68]~[70], [73], [74], [78]~[80], [82], [84], [85], [93], [97], [98], [100], [102], [109], [111], [117], [119]~[121] | 25 |
| (T3) 計測結果の活用 | [29], [46], [49], [50], [52], [53], [58], [62], [87], [94], [99], [104], [106], [113] | 14 |
| (T4) 計測ルールの変更 | [31], [35], [38], [43], [60], [61], [64], [72], [75], [76], [86], [96], [107], [114] | 14 |
| (T5) 計測が難しいとされている対象への適用 | [34], [59], [67], [71], [89], [91], [110], [115], [116] | 9 |
| (T6) その他 | [83], [88], [103], [112], [122] | 5 |

(注1): 論文 [118] は (T1) と (T6) に関連していたが、研究の主目的が「IFPUG 法と COSMIC 法の比較を行い、その利点と欠点の評価」であったため (T1) として分類した。

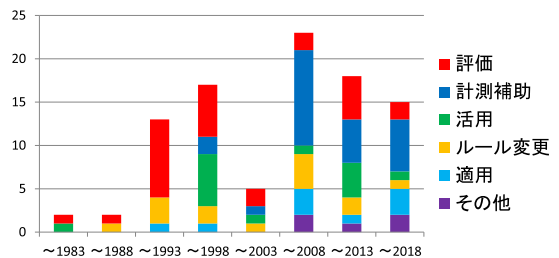


図 2 出版された年代ごとの調査対象論文数

Fig. 2 Number of researched articles by publication age.

研究数が落ち込んでいる。これは 1989 年から 1998 年の 10 年間で FP の利点や活用方法についての研究が多数行われ、それらについての知見が周知されたためだと考えられる。また、2004 年以降に研究数が増加しており、特に (T2) についての研究が多くなっている。(T2) の研究が増加した理由としては、いわゆる設計支援ツールの普及が進んだことが考えられる。従来紙ベースで作成されていた仕様書が電子ファイルとして扱えるようになり、保存されているファイルを直接解析することが可能になった。例えば、UML (Unified Modeling Language) [3] 図が、2003 年に提案された XMI (XML Metadata Interchange) [20] を用いて記述され電子的に扱う事が可能となった事から UML 図を解析して FP を自動計測する研究が行われている [68], [73], [78], [82].

4.1.3 RQ1 への回答

以上より、「(RQ1) FP 研究の動向はどのようなものか」に対しては、前半 20 年においては FP の利点や活用方法に関する研究が、後半 20 年においては FP の計測補助に関する研究が、それぞれ中心に行われていたと回答する。

4.2 RQ2

4.2.1 研究評価の分類

FP についての研究成果が現場で適用可能かどうかを調査するために、各研究で用いられている評価を Zelkowitz の評価モデルを用いて分類し、分析した。分類した結果を表 3 に示す。提案が行われるのみでプロジェクトへの適用などによる評価が行われていない論文が含まれていたため、“No Experimentation” という分類を追加している。

以降、各大分類ごとに評価方法や現場での適用可能

性について述べる。

(1) Observational Methods

この分類に属する研究では、主に提案手法を過去プロジェクト等に適用し従来の FP 法の結果と比較する事でその有用性を評価している。自動計測手法の研究の場合は、手動計測された FP と比較する事で提案手法が従来手法の代替となるかを評価し、計測ルールの変更の場合は変更提案されたルールで計測された FP に基づく見積り工数、従来の FP に基づく見積り工数、実工数を比較することで提案手法が従来手法より良い計測精度であるかどうかを評価している。

研究成果の開発現場での適用可能性を考える上で「研究成果が十分に評価されているか」は重要な要素である。例えば研究成果が少数のプロジェクトでのみ評価されていた場合、その妥当性は低いと考えられる。そこで、本分類の研究で評価に用いたプロジェクト数等を調べた。各研究でのデータの出典及びプロジェクト数をそれぞれ表 4 及び表 5 に示す。54 本中 46 本の研究が企業のプロジェクトによって評価されているため、実際の開発現場に即した環境で評価は行われているように見える。

一方で、必ずしも十分な数のプロジェクトに適用して評価がされているとは限らない。例えば、46 本中 17 本において評価のために用いられたプロジェクト数が 5 件以下であった。また、プロジェクトが単一の企業から得られたものである研究が 46 本中 21 本において存在している。具体的な基準は存在していないが、評価のためのプロジェクト数が少ない場合は、統計的な評価を実施することは難しい。また、単一の企業の

表 3 Zelkowitz の評価モデル
Table 3 Evaluation model of Zelkowitz.

| 大分類 | 小分類 | 論文数 |
|-----------------------|-----------------------------------|-----|
| Observational Methods | Project monitoring | 4 |
| | Case study | 5 |
| | Assertion | 3 |
| | Field study | 42 |
| Historical Methods | Literature search | 3 |
| | Study of Legacy | 18 |
| | Study of Lessons-learned | 6 |
| | Static analysis | 1 |
| Controlled Methods | Replicated Experiment | 0 |
| | Synthetic Environment Experiments | 4 |
| | Dynamic analysis | 0 |
| | Simulation | 2 |
| No Experimentation | | 7 |

表 4 Observational Methods における評価データの出典
Table 4 Source of verification data in Observational Methods.

| データの出典 | 論文数 |
|--------|-----|
| 企業 | 46 |
| 大学 | 5 |
| ISBSG | 3 |

表 5 評価に用いた企業プロジェクトの数
Table 5 Number of corporate projects used for verification.

| プロジェクト数 | 論文数 |
|---------|-----|
| ~5 | 17 |
| 6~50 | 23 |
| 51~100 | 3 |
| 101~ | 2 |
| 不明 | 1 |

データから得られた結論は、必ずしも他の企業においても成り立つとは限らない。

研究成果の開発現場での適用可能性を考える上では、研究成果が適用可能である対象のコンテキスト [21] についての情報が重要である。コンテキストとは、プロジェクトにおける開発言語や開発規模、業種等の特徴を表すものである。多くの研究において、研究成果は一定のコンテキストをもつプロジェクトで適用が可能である場合が多い。開発現場において研究成果が適用されたプロジェクトと類似するコンテキストをもったプロジェクトが扱われていた場合、そのプロジェクトへの研究成果の適用可能性は高いと考えられる。したがって、研究成果がどのようなコンテキストの上で成り立つかが詳細に記載されているかどうか、開発現場に研究成果を導入するための有効な判断材料となる。

そこで、Observational Methods に分類された論文において、評価対象のプロジェクトのコンテキストが十分記載されているかどうかを調査した。今回は独立行政法人情報処理推進機構 (IPA) の社会基盤センターが発行しているデータ白書 [16] に記載されているプロジェクトのプロファイルデータの項目をコンテキスト情報として選択し、それらについての情報が論文に記載されているかどうかを調べた。

コンテキストとして扱われるデータ項目を表 6 に示す。対象論文において、コンテキスト項目が全て記載されている論文は存在しなかった。また、項目を「開発プロジェクトの種別」、「業種」、「開発言語」に絞った場合でもそれらの項目が全て記載されている論文は 54 本中 12 本しか存在せず、大半の論文において適用可能なプロジェクトのコンテキストが記載されていなかった。そのため、研究成果が開発現場へ適用可能かを判断するための情報が不足している状況にあり、改善が必要であると判断できる。

表 6 プロジェクトのコンテキスト
Table 6 Project context.

| コンテキスト | 概要 |
|-------------|---|
| 開発プロジェクトの種別 | 新規開発, 改修・保守, 拡張 |
| アーキテクチャ | イントラネット/インターネット, 2 階層クライアント/サーバ, 3 階層クライアント/サーバ |
| 業種 | 金融・保険業, 情報通信業, 製造業, 公務など |
| 開発言語 | Java, VB, C, COBOL, C++ など |
| プラットフォーム | Windows 系, Unix 系 |
| その他 | 工数, 規模, 工期 |

(2) Historical Methods

ここでは、過去の文献やプロジェクトを分析する事によって知見を得る研究が行われていた。前述の Observational Methods や Controlled Methods のように研究成果の評価は行われておらず、その妥当性は知見を得るために用いたプロジェクトの出典や数に比例すると考えられる。各研究に用いられたデータの出典及びプロジェクト数をそれぞれ表 7 及び表 8 に示す。

小分類の違いが「評価対象となるプロジェクトの数や規模」でしかなく提案手法の評価という点で同一であった Observational Methods と比べ、Historical Methods は小分類によって行われている内容が大きく異なるため各小分類ごとに分析する。

- Literature search : 研究論文を元に行われた研究は 3 件存在した。内 1 件は FP の問題点とその改善についての SLR であり、残り 2 件は既存の研究で得られた FP の利点等をまとめた論文であった。この小分類の研究は既存研究を元としているため、既存研究に妥当性がない場合以外は得られた知見に妥当性があると言える。

- Study of Legacy : この分類には、過去のプロジェクトを分析する事で FP の計測要素と工数の関係性を調査する論文などが含まれている。過去プロジェクトから知見を得る以上、得られた知見の妥当性は知見を得るために用いられたプロジェクトの出典及び件数に依存すると考えられる。これは、例えば大学のプログラミング演習での開発から得られた知見は企業の開発現場にはそぐわない可能性や、少数のプロジェ

表 7 Historical Methods におけるデータの出典
Table 7 Source of data in Historical Methods.

| データの出典 | 論文数 |
|--------|-----|
| 企業 | 20 |
| ISBSG | 4 |
| 研究論文 | 3 |
| 大学 | 1 |

表 8 Historical Methods におけるプロジェクト数
Table 8 Number of projects in Historical Methods.

| プロジェクト数 | 論文数 |
|---------|-----|
| ~10 | 7 |
| 11~30 | 4 |
| 31~100 | 3 |
| 101~ | 7 |
| 研究論文 | 3 |
| アンケート | 3 |

クトから得られた知見には一般性がない可能性があるためである。この小分類における研究に用いられたデータの出典及びプロジェクト数をそれぞれ表 9 及び表 10 に示す。過去のプロジェクトは主に企業のプロジェクト及び様々な国の企業のプロジェクトが集められたデータセットである ISBSG [14] から得られており、出典に基づく妥当性は高いと言える。一方、知見を得るために用いられたプロジェクト数としては、100 を超えるプロジェクトを用いた論文が 6 本存在するものの、10 件以下という少数のプロジェクトを用いて分析を行った論文も存在しているため、この小分類において開発現場への適用可能性が高いとは一概に言えない。

- Study of Lessons-learned: この分類においては、プロジェクトを実行した際の情報や開発者へのインタビュー等により、FP についての教訓をまとめる論文が含まれている。基本的に、単一のプロジェクトを開発した際の得られた情報やその際関わっていた開発者へのインタビューを元にしてしているため得られた教訓は企業が取り扱う業種や人員に依存している可能性が高く、そのため妥当性は低いものの、似た業種、似た環境の開発現場であれば研究成果の適用可能性は高いと考えられる。

(3) Controlled Methods

ここでは、主に提案手法をあるプロジェクトに適用し従来の FP 法の結果等と比較する事でその有用性を評価していた。同じく提案手法をプロジェクトに適用し評価する Observational Methods との違いとしては、その厳密な評価方法が挙げられる。Observational Methods では主に研究者自らが提案手法による計測を行っていたため、従来手法での結果との差に人的要

因が含まれている可能性があった。それに対し、Controlled Methods は評価のために多くの人々を用いる事で人的要因を消去するなど、より厳密に提案手法を評価している。しかし、複数の人々に評価させる事によって人的要因をなくし、より厳密な評価を行う手法である Synthetic Environment Experiments はその必要な人員数から評価コストが高く実行が難しい。そのため、95 本の研究の中でも 4 件のみしか行われていない。また、そのコストから Observational Methods のように多くのプロジェクトでの評価が困難であり、4 件の研究においても評価に用いられたプロジェクト数は 3 件が 1 プロジェクト、1 件が 2 プロジェクトと極めて少数となっている。Observational Methods と Controlled Methods を比べた場合、前者は「提案手法はどんなプロジェクトでも適用が可能か」を評価する側面が強く、後者は「提案手法は誰が行っても同じ結果が得られるか」を評価する側面が強い。そのため、研究によってどちらが評価方法として適しているかを考慮する必要がある。また、Simulation については、提案手法が適用される場面を実際にシミュレートし評価する事が可能であるため、Controlled Methods の中でも特に妥当性が高い評価手法だと言える。しかしながら、該当する研究はどちらも 1990 年代の物と古く、当時と比べて現在のシステムは非常に複雑化しているため、現状におけるシミュレーションによるシステムの再現及び提案手法の評価は難しいと考えられる。

4.2.2 RQ2 への回答

以上より、「(RQ2) 研究成果は現場での適用が可能か」に対しては、各論文で採用されている三つの評価手法により結論が異なる。Observational Methods が用いられている研究については、研究成果の開発現場への適用可能性を判断するための情報が不足しているため、改善が必要である。Historical Methods が用いられている研究については、Literature search と Study of Lessons-learned では適用可能性が高く、Study of Legac では必ずしも適用可能性が高いといえない。Controlled Methods についても、必ずしも適用可能性が高いといえない。

4.3 RQ3

4.3.1 開発現場のニーズの収集

開発現場のニーズと実際に行われている研究にギャップが存在するかを確認するために、本論文では、JF-PUG (日本ファンクションポイントユーザ会) の FP 活用研究会で得られた FP に対する開発現場のニーズ

表 9 Study of Legacy におけるデータの出典
Table 9 Source of data in Study of Legacy.

| データの出典 | 論文数 |
|--------|-----|
| 企業 | 15 |
| ISBSG | 4 |
| 大学 | 1 |

表 10 Study of Legacy におけるプロジェクト数
Table 10 Number of projects in Study of Legacy.

| プロジェクト数 | 論文数 |
|---------|-----|
| ～10 | 6 |
| 11～30 | 4 |
| 31～100 | 3 |
| 101～ | 7 |

と既存研究を照らし合わせ、各ニーズに対応する研究があるかどうかを調べた。

得られたニーズは、2018年度のFP活用研究会の活動の中で、研究会参加者(10名程度、主にベンダ企業の所属)が自社においてFPを導入する上で問題とされている意見の中で、主に研究として扱うべきものを収集したものである。ニーズとして抽出された際には、ニーズ内での重要度は考慮されていない。

JFPUGはIFPUG(International Function Point User Group)の日本支部であり、日本国内におけるFPの理解と普及、その利用法の検討及びソフトウェア計測の検討を主な目的とし、IFPUGで検討されている様々な課題を共有している。また、FP活用研究会は現場におけるFP活用の課題に対する解決策を探求し、定量データを利用する様々な活動の改善につなげることを目的としていることより、この研究会で得られたニーズは現場のニーズを反映した妥当なものであると考えている。

4.3.2 RQ3への回答

得られたニーズと対応する論文の本数を表11に示す。

六つのニーズの内、「保守案件への適用が難しい」、「FPを用いた生産性評価モデルが確立されていない」、「FPで計測できない要素への対応」、「発注側企業にFPが普及していない」、「FPのよい教育手法がない」という五つのニーズにおいて対応する研究数は「汎用的な自動計測ツールがない」というニーズに比べて少ない。特に、「FPのよい教育手法がない」というニーズに対応する研究は1本も存在しなかった。この結果から、開発現場に対するニーズに対して、必ずしも十分に研究が実施されているとは言えない。

表11 開発現場のニーズと対応する論文の本数

Table 11 Number of papers corresponding to the needs of the development site.

| 開発現場のニーズ | 対応する研究 | 論文数 |
|-------------------------|--|-----|
| 汎用的な自動計測ツールがない | [55], [57], [65], [68], [69], [73], [74], [78], [79], [85], [98], [100], [117], [119], [120] | 15 |
| 保守案件への適用が難しい | [59], [115], [116] | 3 |
| FPを用いた生産性評価モデルが確立されていない | [29], [62] | 2 |
| FPで計測できない要素への対応 | [106] | 1 |
| 発注側企業にFPが普及していない | [113] | 1 |
| FPのよい教育手法がない | - | 0 |

なお、「汎用的な自動計測ツールがない」ということに関しては、他のニーズに比べて多くの研究が行われてきており、計測ツールの開発や評価が報告されているが、ある組織のコンテキストに依存しすぎているツールや現場導入において非常に手間がかかるものが多いため、すぐに現場へ導入可能なものがないという点で研究の余地があると考えられる。

以上より、「(RQ3) 開発現場のニーズと研究にギャップは存在するか」に対しては、ギャップが存在すると回答する。

5. 考 察

ここではRQへの回答を元に、FP研究における今後の課題を考察する。

5.1 コンテキスト情報の定義

Observational Methodsにおいて、多くの研究において研究成果が適用可能なプロジェクトのコンテキストについての情報が記載されておらず、そのため研究成果の現場への適用可能性を判断することが難しい現状が存在する。これを解決するためには今後の研究において研究成果がどのようなコンテキストをもったプロジェクトに対して適用可能であるかを詳細に記述していく必要がある。また、現状においては開発現場への研究成果の適用を考えるうえでどのようなコンテキスト情報が有用であるかが不明である。4.2.1では、開発現場への適用という観点から我が国でよく知られているIPAのデータ白書を用いたが、他にもコンテキスト情報として利用されているものは複数存在する[7], [14]。これらを元にして、現場への適用可能性を判断するために必要なコンテキスト情報を定義することは今後の重要な課題である。

5.2 教育手法の開発

FP導入にはFPや見積りに関する学習が必要であるが、FPの教育手法に関する研究は今回対象とした論文の中では行われていなかった。効果的なFP教育法の開発は、FP利活用の裾野を広げる意味でも重要な課題である。

FP計測において重要な点は“計測ルールの理解”と“要求仕様の補完”である。計測ルールを誤って理解していた場合計測結果には誤差が生じるため、計測ルールの正しい理解は不可欠である。そのため、教育方法としては計測ルールを正しく理解させるための多くの例題が必要である。また、要求仕様などからFPを計測する場合、仕様に記載されている機能の抽象度が高

いため、FP 計測時に計測者が要求仕様の補完をする必要がある。例えば、要求仕様上では一つの機能として記載されているものが、複数の小さい機能の組み合わせで実現されることがある。その場合 FP を計測する際も複数の機能として計測する必要がある。その際、経験を積んだ計測者は業務知識等を元に自ら仕様を補完して計測を行えるが、経験が浅い者は補完を行えない。これを解決するためには、実際の設計・実装経験や FP の計測経験を積むことが必要ではあるが、教育において FP 計測時に実際に行われた仕様の補完の実績を蓄積し参照する事で要求仕様の補完方法を学習することも有効である。そこで、“どのような例題を用意すれば FP の計測ルールを正確に理解させる事が可能であるか”、“要求仕様の補完について学習するためにはどのような情報を蓄積し共有すればよいか”を特定する研究が考えられる。

5.3 FP で計測できない要素への対応

FP は機能規模を計測する手法であるため、工数見積りに利用する場合は、FP を非機能要件等も考慮する必要がある。IFPUG 法に関しては、FP 提案当初、非機能要件は調整係数として考慮され、非機能要件の評価結果に応じて $\pm 35\%$ の範囲で調整がされていたが、2009 年に ISO 標準に認定されて以降、非機能要件の調整は IFPUG 法としては含まれなくなった。近年では非機能要求に対する重要性が増しており [15]、上記の問題への対策が求められている。

このニーズに対応した既存研究として、FP とそれ以外の要素を用いた工数見積りについての研究 [106] が存在する。この研究では、ニューラルネットワークの入力として開発規模である FP のほかにプロジェクトの特徴やデータベースの種類、開発言語などといった非機能要件を用意し、工数見積りにおける精度を比較した。“FP が入力に含まれている見積りはそうでないものと比べて高精度である”、“FP とデータベースの種類を入力とした見積りが最も高精度である”という結果が得られたため、開発規模が工数に最も強い影響を及ぼす事、そして開発規模による工数見積りは非機能要件を組み合わせる事で見積り精度が向上される事が分かった。今後も、FP と組み合わせる事で工数見積り精度を向上させるような非機能要件を特定する研究が必要である。

6. む す び

本研究では、FP 研究における研究動向の調査を目的とし、過去 40 年間に出版された FP に関する論文を対象とした系統的文献レビューを実施した。95 本の論文を対象に研究内容の整理と研究動向の分析を行った結果、1979 年の提案後 20 年間では“他の開発規模尺度との比較等による FP 法の利点や欠点の評価”と“計測された FP の工数見積りや生産性評価への活用”が多く研究され、直近 20 年においては“計測の手間の削減などのための FP の自動計測手法”についての研究が主流である事が分かった。また、研究成果の開発現場での適用可能性や開発現場のニーズと研究とのギャップについて調査を行った。

今後の課題として、5. で述べた、“コンテキスト情報の定義”、“教育手法の開発”、“FP で計測できない要素への対応”が、挙げられる。また、文献 [112] も FP に関する系統的文献レビューが実施されているが、その目的は「FP 法の計測精度の改善」の研究調査であり、対象は 2002 年から 2013 年までに発表された 18 本の論文である。文献 [112] は本論文の分類では T4 (計測ルールの変更) に対応しており、より詳細な分析を行った上で、FP 計測における基本機能要素の複雑度の重みづけ方法や非機能要件を用いた FP 値の調整についての改善について述べられている。個々のトピックについてより詳細な分析をすることも今後の課題の一つである。

謝辞 本研究を進めるにあたって、貴重なご助言を頂きました日本ファンクションポイントユーザ会 FP 活用研究会の皆様へ深く感謝申し上げます。

文 献

- [1] ACM, “ACM Digital library,” <http://dl.acm.org/>, 参照 March 18, 2019.
- [2] A.J. Albrecht, Function point analysis, Encyclopedia of Software Engineering, vol.1, pp.518–524, Addison-Wesley Professional, 1994.
- [3] S.W. Ambler, The object primer: Agile model-driven development with UML 2.0. Cambridge University Press, 2004.
- [4] B. Boehm, C. Abts, and S. Chulani, “Software development cost estimation approaches - A survey,” Annals of Software Engineering, vol.10, no.1-4, pp.177–205, Springer, 2000.
- [5] P. Brereton, B. Kitchenham, D. Budgen, M. Turner, and M. Khalil, “Lessons from applying the systematic literature review process within the software engineering domain,” J. Syst. Softw., vol.80, no.4, pp.571–583, 2007.
- [6] CORE, “Computing Research & Education; CORE Rankings Portal”. <http://www.core.edu.au/conference-portal>, 参照 March 18, 2019.

- [7] R.L. Feldmann, J. Munch, and S. Vorwieger, "Towards goal-oriented organizational learning: representing and maintaining knowledge in an experience base," Proc. Tenth International Conference on Software Engineering and Knowledge Engineering, pp.236–245, 1998.
- [8] R.L. Glass, Facts and fallacies of software engineering, Addison-Wesley Professional, 2002.
- [9] Google, "Google Scholar," <https://scholar.google.co.jp/>, 参照 March 18, 2019.
- [10] J.E. Hirsch, "An index to quantify an individual's scientific research output," Proc. National academy of Sciences, vol.102, no.46, pp.16569–16572, 2005.
- [11] V. Ivanov, A. Rogers, G. Succi, J. Yi, and V. Zorin, "What do software engineers care about? gaps between research and practice," Proc. 2017 11th Joint Meeting on Foundations of Software Engineering, pp.890–895, 2017.
- [12] IEEE, "IEEE xplore," <http://ieeexplore.ieee.org/Xplore/home.jsp>, 参照 March 18, 2019.
- [13] IFPUG, "Function Point Counting Practices Manual, Release 4.3.1," International Function Point Users Group, 2010.
- [14] ISBSG, <https://www.isbsg.org/data-collection-questionnaires>, 参照 March 18, 2019.
- [15] 情報処理通信機構, "非機能要求グレード 2018 利用ガイド," <https://www.ipa.go.jp/sec/softwareengineering/std/ent03-b.html>, 参照 Aug. 13, 2019.
- [16] 情報処理通信機構, "ソフトウェア開発データ白書 2016-2017," <https://www.ipa.go.jp/files/000057877.pdf>, 参照 March 18, 2019.
- [17] B. Kitchenham, "Procedures for performing systematic reviews," Keele, UK, Keele University, vol.33, no.2004, pp.1–26, 2004.
- [18] B. Kitchenham and P. Brereton, "A systematic review of systematic review process research in software engineering," Information and Software Technology, vol.55, no.12, pp.2049–2075, 2013.
- [19] B. Kitchenham, P. Brereton, Z. Li, D. Budgen, and A. Burn, "Repeatability of systematic literature reviews," Evaluation & Assessment in Software Engineering, pp.46–55, 2011.
- [20] Object Management Group, "ABOUT THE XML METADATA INTERCHANGE SPECIFICATION VERSION 2.0," <https://www.omg.org/spec/XMI/2.0/About-XMI/>, 参照 March 18, 2019.
- [21] P. Oman and S.L. Pfleeger, Applying Software Metrics, John Wiley & Sons, 1996.
- [22] Science Direct, "Science Direct," <http://www.sciencedirect.com/>, 参照 March 18, 2019.
- [23] Scopus, "Scopus," <https://www.scopus.com/>, 参照 March 18, 2019.
- [24] 総務省行政管理局. 政府情報システムの整備及び管理に関する標準ガイドライン実務手引書. 政府情報システムの整備及び管理に関する標準ガイドライン実務手引書, p.122, 内閣官房情報通信技術 (IT) 総合戦略室, 2015.
- [25] C.R. Symons, Software sizing and estimating: Mk II FPA (function point analysis), John Wiley & Sons, 1991.
- [26] 山口正明 (著), 調重俊 (監修), ファンクションポイント COSMIC-FFP 法実践ガイド—組込み系・リアルタイム系に最適なソフトウェア規模・工数の見積り方法, 日科技連出版社, 2007.
- [27] M.V. Zelkowitz and D.R. Wallace, "Experimental models for validating technology," Computer, vol.31, no.5, pp.23–31, 1998.
- [28] A.J. Albrecht and J.E. Gaffney, "Software function, source lines of code, and development effort prediction: a software science validation," IEEE Trans. Softw. Eng., vol.6, pp.639–648, 1983.
- [29] C.A. Behrens, "Measuring the productivity of computer systems development activities with function points," IEEE Trans. Softw. Eng., vol.6, pp.648–652, 1983.
- [30] C.F. Kemerer, "An empirical validation of software cost estimation models," Commun. ACM, vol.30, no.5, pp.416–429, 1987.
- [31] C.R. Symons, "Function point analysis: Difficulties and improvements," IEEE Trans. Softw. Eng., vol.14, no.1, pp.2–11, 1988.
- [32] J.M. Verner, G. Tate, B. Jackson, and R.G. Hayward, "Technology dependence in function point analysis: A case study and critical review," 11th International Conference on Software Engineering, pp.375–382, 1989.
- [33] G.C. Low and D.R. Jeffery, "Function points in the estimation and evaluation of the software process," IEEE Trans. Softw. Eng., vol.1, pp.64–71, 1990.
- [34] D.J. Reifer, "Asset-R: A function point sizing tool for scientific and real-time systems," J. Syst. Softw., vol.11, no.3, pp.159–171, 1990.
- [35] L.G. Guerra and H.S. Lobos, "An empirical validation of a software cost model," IFAC Proceedings, vol.23, no.5, pp.57–62, 1990.
- [36] F.J. Heemstra and R.J. Kusters, "Function point analysis: Evaluation of a software cost estimation model," European Journal of Information Systems, vol.1, no.4, pp.229–237, 1991.
- [37] M. van Genuchten and H. Koolen, "On the use of software cost models," Information & Management, vol.21, no.1, pp.37–44, 1991.
- [38] D.B. Bock and R. Klepper, "FP-S: A simplified function point counting method," J. Syst. Softw., vol.18, no.3, pp.245–254, 1992.
- [39] C.F. Kemerer, F. Chris, and S.P. Benjamin, "Improving the reliability of function point measurement: an empirical study," IEEE Trans. Softw. Eng., vol.18, no.11, pp.1011–1024, 1992.
- [40] R. Rask, P. Laamanen, and K. Lyytinen, "A comparison of Albrecht's function point and Symons' Mark II

- metrics,” International Conference on Software and Information Systems, pp.207–221, 1992.
- [41] R. Betteridge, “Successful experience of using function points to estimate project costs early in the lifecycle,” *Information and Software Technology*, vol.34, no.10, pp.655–658, 1992.
- [42] B. Kitchenham and K. Kansala, “Inter-item correlations among function points,” *Proc. First International Software Metrics Symposium*, pp.11–14, 1993.
- [43] D.R. Jeffery, G.C. Low, and M. Barnes, “A comparison of function point counting techniques,” *IEEE Trans. Softw. Eng.*, vol.19, no.5, pp.529–532, 1993.
- [44] R. Rask, P. Laamanen, and K. Lyytinen, “Simulation and comparison of Albrecht’s function point and DeMarco’s function bang metrics in a CASE environment,” *IEEE Trans. Softw. Eng.*, vol.19, no.7, pp.661–671, 1993.
- [45] R. Hughes, “An exploration of the problems of using function points mark II,” *Information Systems Journal*, vol.4, no.3, pp.169–183, 1994.
- [46] J.E. Matson, B.E. Barrett, and J.M. Mellichamp, “Software development cost estimation using function points,” *IEEE Trans. Softw. Eng.*, vol.20, no.4, pp.275–287, 1994.
- [47] A. Abran and P.N. Robillard, “Function points: a study of their measurement processes and scale transformations,” *J. Syst. Softw.*, vol.25, no.2, pp.171–184, 1994.
- [48] B. Kitchenham, S.L. Pfleeger, and N. Fenton, “Towards a framework for software measurement validation,” *IEEE Trans. Softw. Eng.*, vol.21, no.12, pp.929–944, 1995.
- [49] C. Jones, “Backfiring: Converting lines of code to function points,” *Computer*, vol.28, no.11, pp.87–88, 1995.
- [50] C. Jones, “Determining software schedules,” *Computer*, vol.28, no.2, pp.73–75, 1995.
- [51] A. Abran and P.N. Robillard, “Function points analysis: An empirical study of its measurement processes,” *IEEE Trans. Softw. Eng.*, vol.22, no.12, pp.895–910, 1996.
- [52] G.H. Subramanian and G.E. Zarnich, “An examination of some software development effort and productivity determinants in ICASE tool projects,” *J. Management Information Systems*, vol.12, no.4, pp.143–160, 1996.
- [53] G.R. Finnie, G.E. Wittig, and J.M. Desharnais, “A comparison of software effort estimation techniques: using function points with neural networks, case-based reasoning and regression models,” *J. Syst. Softw.*, vol.39, no.3, pp.281–289, 1997.
- [54] S. Furey, “Why we should use function points [software metrics],” *IEEE Softw.*, vol.14, no.2, p.28, 1997.
- [55] P. Shoval and O. Feldman, “A combination of the Mk-II function points software estimation method with the ADISSA methodology for systems analysis and design,” *Information and Software Technology*, vol.39, pp.855–865, 1997.
- [56] J.J. Dolado, “A study of the relationships among Albrecht and Mark II function points, lines of code 4GL and effort,” *J. Syst. Softw.*, vol.37, no.2, pp.161–173, 1997.
- [57] G. Antonioli, F. Calzolari, L. Cristoforetti, R. Fiutem, and G. Caldiera, “Adapting function points to object oriented information systems,” *International Conference on Advanced Information Systems Engineering CAiSE 1998: Advanced Information Systems Engineering*, pp.59–76, 1998.
- [58] F. Fujiwara, T. Goto, and S. Araki, “Examples of applying software estimate tool,” *Proc. 20th International Conference on Software Engineering (ICSE ’98)*, pp.469–472, 1998.
- [59] R.A. Fleck, Jr, “Managing programmer resources in a maintenance environment with function points,” *Industrial Management & Data Systems*, vol.98, no.2, pp.63–70, 1998.
- [60] C. Yau and H.L. Tsoi, “Modelling the probabilistic behaviour of function point analysis,” *Information and Software Technology*, vol.40, no.2, pp.59–68, 1998.
- [61] G. Horgan, S. Khaddaj, and P. Forte, “Construction of an FPA-type metric for early lifecycle estimation,” *Information and Software Technology*, vol.40, no.8, pp.409–415, 1998.
- [62] H.S. Bok and K.S. Raman, “Software engineering productivity measurement using function points: a case study,” *J. Inf. Technol.*, vol.15, no.1, pp.79–90, 2000.
- [63] C.J. Lokan, “An empirical analysis of function point adjustment factors,” *Information and Software Technology*, vol.42, no.9, pp.649–659, 2000.
- [64] G. Orr and T.E. Reeves, “Function point counting: One program’s experience,” *J. Syst. Softw.*, vol.53, no.3, pp.239–244, 2000.
- [65] S. Kusumoto, M. Imagawa, S. Morimoto, K. Matsusita, and M. Tsuda, “Function point measurement from Java programs,” *Proc. 24th International Conference on Software Engineering (ICSE ’02)*, pp.576–582, 2002.
- [66] B. Kitchenham, S.L. Pfleeger, B. McColl, and S. Eaganc, “An empirical study of maintenance and development estimation accuracy,” *J. Syst. Softw.*, vol.64, no.1, pp.57–77, 2002.
- [67] A.T. Calazans, K.M. de Oliveira, and R.R. Santos, “Adapting function point analysis to estimate data mart size,” *Proc. 10th International Software Metrics Symposium*, pp.300–311, 2004.
- [68] G. Cantone, D. Pace, and G. Calavaro, “Applying function point to unified modeling language: Conversion model and pilot study,” *Proc. 10th International*

- Software Metrics Symposium, pp.280–291, 2004.
- [69] E. Lamma, P. Mello, and F. Riguzzi, “A system for measuring function points from an ER-DFD specification,” *The Computer Journal*, vol.47, no.3, pp.358–372, 2004.
- [70] P. Umbers and G. Miles, “Resource estimation for Web applications,” *Proc. 10th International Software Metrics Symposium*, pp.370–381, 2004.
- [71] M. Sanchez-Segura, J.J. Cuadrado, A.M. Moreno, A. de Amescua, A. de Antonio, and O. Marbana, “Virtual reality systems estimation vs.traditional systems estimation,” *J. Syst. Softw.*, vol.72, no.2, pp.187–194, 2004.
- [72] M.A. Al-Hajri, A.A.A. Ghani, M.N. Sulaiman, and M.H. Selamat, “Modification of standard function point complexity weights system,” *J. Syst. Softw.*, vol.74, no.2, pp.195–206, 2005.
- [73] H. Diab, F. Koukane, M. Frappier, and R. St-Denis, “ μ CRose: Automated measurement of COSMIC-FFP for rational rose realtime,” *Information and Software Technology*, vol.47, no.3, pp.151–166, 2005.
- [74] V. Harput, H. Kaindl, and S. Kramer, “Extending function point analysis to object-oriented requirements specifications,” *Proc. 11th IEEE International Software Metrics Symposium*, pp.361–370, 2005.
- [75] T. Kralj, I. Roxman, M. Hericko, and A. Zivkovic, “Improved standard FPA method-resolving problems with upper boundaries in the rating complexity process,” *J. Syst. Softw.*, vol.77, no.2, pp.81–90, 2005.
- [76] L. Santillo, M. Conte, and R. Meli, “Early& Quick function point: sizing more with less,” *Proceedings of the 11th International Software Metrics Symposium*, pp.378–383, 2005.
- [77] S. Sheetz, L. Wallace, and D. Henderson, “Understanding manager and developer perceptions of the relative advantage, compatibility, and complexity of function points and source lines of code,” *11th Americas Conference on Information Systems*, pp.3017–3026, USA, 2005.
- [78] A. Zivkovic, I. Rozman, and M. Hericko, “Automated software size estimation based on function points using UML models,” *Information and Software Technology*, vol.47, no.13, pp.881–890, 2005.
- [79] M. Hericko, I. Rozman, and A. Zivkovic, “A formal representation of functional size measurement methods,” *J. Syst. Softw.*, vol.79, no.9, pp.1341–1358, 2006.
- [80] S. Abrahao and G. Poels, “Experimental evaluation of an object-oriented function point measurement procedure,” *Information and Software Technology*, vol.49, no.4, pp.366–380, 2007.
- [81] C. Gencel and O. Demirors, “Conceptual differences among functional size measurement methods,” *First International Symposium on Empirical Software Engineering and Measurement*, pp.305–313, 2007.
- [82] V. Del Bianco, C. Gentile, and L. Lavazza, “An evaluation of function point counting based on measurement-oriented models,” *Proc. 12th International Conference on Evaluation and Assessment in Software Engineering*, pp.11–20, 2008.
- [83] C. Gallego, L. Buglione, R.J. Rejas-Muslera, and F. Machado, “IFPUG-cosmic statistical conversion,” *Proc. 34th Euromicro Conference Software Engineering and Advanced Applications (SEAA '08)*, pp.427–432, 2008.
- [84] L.A. Lavazza, V. Del Bianco, and C. Garavaglia, “Model-based functional size measurement,” *Proc. Second ACM-IEEE International Symposium on Empirical Software Engineering and Measurement*, pp.100–109, 2008.
- [85] S. Abrahao and E. Insfran, “A metamodeling approach to estimate software size from requirements specifications,” *34th Euromicro Conference Software Engineering and Advanced Applications*, pp.465–475, 2008.
- [86] W. Xia, L.F. Capretz, D. Ho, and F. Ahmed, “A new calibration for Function Point complexity weights,” *Information and Software Technology*, vol.50, no.7-8, pp.670–683, 2008.
- [87] H. Park and S. Baek, “An empirical validation of a neural network model for software effort estimation,” *Expert Systems with Applications*, vol.35, no.3, pp.929–937, 2008.
- [88] C. Gallego, F. Machado-Piriz, and J. Aroba-Paez, “On the conversion between IFPUG and COSMIC software functional size units: A theoretical and empirical study,” *J. Syst. Softw.*, vol.81, no.5, pp.661–672, 2008 .
- [89] M. Hericko and A. Zivkovic, “The size and effort estimates in iterative development,” *Information and Software Technology*, vol.50, no.7-8, pp.772–781, 2008.
- [90] P. Kampstra and C. Verhoef, “Reliability of function point counts,” *Commun. ACM*, pp.85–97, 2009.
- [91] L. Lavazza and C. Garavaglia, “Using function points to measure and estimate real-time and embedded software: Experiences and guidelines,” *Proc. 3rd International Symposium on Empirical Software Engineering and Measurement (ESEM '09)*, pp.100–110, 2009.
- [92] O. Demirors and C. Gencel, “Conceptual association of functional size measurement methods,” *IEEE Softw.*, vol.26, no.3, pp.71–78, 2009.
- [93] V. Del Bianco and L. Lavazza, “Applying the COSMIC functional size measurement method to problem frames,” *Proc. 14th IEEE International Conference on Engineering of Complex Computer Systems*, pp.282–290, 2009.
- [94] S. Bagheri, K. Ratakonda, and R. Mohan, “Variable Productivity Adjustment Estimation for Function

- Point Project Delivery,” IEEE International Conference on Data Mining Workshops, pp.557–561, 2010.
- [95] K. Lind and R. Heldal, “Categorization of real-time software components for code size estimation,” Proc. 2010 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement, article no.26, pp.1–10, 2010.
- [96] L. Lavazza and G. Robiolo, “Introducing the evaluation of complexity in functional size measurement: a UML-based approach,” Proc. 2010 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement, article no.25, pp.1–9 2010.
- [97] S. Abrahao, J. Gomez, and E. Insfran, “Validating a size measure for effort estimation in model-driven Web development,” *Inf. Sci.*, vol.180, no.20, pp.3932–3954, 2010.
- [98] T. Edagawa, T. Akaike, Y. Higo, S. Kusumoto, S. Hanabusa, and T. Shibamoto, “Function point measurement from Web application source code based on screen transitions and database accesses,” *J. Syst. Softw.*, vol.84, no.6, pp.976–984, 2011.
- [99] E.A. El-Sebakhy, “Functional networks as a novel data mining paradigm in forecasting software development efforts,” *Expert Systems with Applications*, vol.38, no.3, pp.2187–2194, 2011.
- [100] K. Lind and R. Heldal, “A model-based and automated approach to size estimation of embedded software components,” *International Conference on Model Driven Engineering Languages and Systems*, pp.334–348, 2011.
- [101] F.G. Wilkie, I.R. McChesney, P. Morrow, C. Tuxworth, and N.G. Lester, “The value of software sizing,” *Information and Software Technology*, vol.53, no.11, pp.1236–1249, 2011.
- [102] S. Choi, S. Park, and V. Sugumaran, “A rule-based approach for estimating software development cost using function point and goal and scenario based requirements,” *Expert Systems with Applications*, vol.39, no.1, pp.406–418, 2012.
- [103] A.Z. Abualkishik, J.M. Desharnais, A. Lhelifi, A.A.A. Ghani, R. Atan, and M.H. Selamat, “An exploratory study on the accuracy of FPA to COSMIC measurement method conversion types,” *Information and Software Technology*, vol.54, no.11, pp.1250–1264, 2012.
- [104] M. Tsunoda, Y. Kamei, K. Toda, M. Nagappan, K. Fushida, and N. Ubayashi, “Revisiting software development effort estimation based on early phase development activities,” Proc. 10th IEEE Working Conference on Mining Software Repositories, pp.429–438, 2013.
- [105] M. Staples, R. Kolanski, G. Klein, C. Lewis, J. Andronick, T. Murray, R. Jeffery, and L. Bass, “Formal specifications better than function points for code sizing,” *International Conference on Software Engineering*, pp.1257–1260, 2013.
- [106] D.R. Pai, K.S. McFall, and G.H. Subramanian, “Software effort estimation using a neural network ensemble,” *J. Computer Information Systems*, vol.53, no.4, pp.49–58, 2013.
- [107] L. Lavazza, S. Morasca, and G. Robiolo, “Towards a simplified definition of function points,” *Information and Software Technology*, vol.55, no.10, pp.1796–1809, 2013.
- [108] C. Quesada-Lopez and M. Jenkins, “Function point structure and applicability validation using the ISBSG dataset: A replicated study,” Proc. 8th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement, p.66, 2014.
- [109] V. Del Bianco, L. Lavazza, G. Liu, S. Morasca, and A.Z. Abualkishik, “Model-based early and rapid estimation of COSMIC functional size-An experimental evaluation,” *Information and Software Technology*, vol.56, no.10, pp.1253–1267, 2014.
- [110] F. Paz, C. Zapata, and J.A. Pow-Sang, “An approach for effort estimation in incremental software development using cosmic function points,” Proc. 8th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement, pp.44:1–44:4, 2014.
- [111] D. Ceke and B. Milasinovic, “Early effort estimation in web application development,” *J. Syst. Softw.*, vol.103, pp.219–237, 2015.
- [112] M. de Freitas Junior, M. Fantinato, and V. Sun, “Improvements to the function point analysis method: A systematic literature review,” *IEEE Trans. Engineering Management*, vol.62, no.4, pp.495–506, 2015.
- [113] H. Huijgens, G. Gousios, and A. van Deursen, “Pricing via functional size—a case study of a company’s portfolio of 77 outsourced projects,” *ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*, pp.182–191, 2015.
- [114] A. Hira, H. Hakim, A. Abran, and H. Ben-Abdallah, “A measurement method for sizing the structure of UML sequence diagrams,” *Information and Software Technology*, vol.59, pp.222–232, 2015.
- [115] A. Hira and B. Boehm, “Function point analysis for software maintenance,” Proc. 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement, pp.48:1–48:6, 2016.
- [116] A. Hira and B. Boehm, “Using software non-functional assessment process to complement function points for software maintenance,” Proc. 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement, pp.50:1–50:6, 2016.
- [117] S. Bagriyanik and A. Karahoca, “Automated COSMIC function point measurement using a requirements engineering ontology,” *Information and Software Technology*, vol.72, pp.189–203, 2016.

- [118] S.D. Martino, F. Ferrucci, C. Gravino, and F. Sarro, "Web effort estimation: function point analysis vs. COSMIC," *Information and Software Technology*, vol.72, pp.90-109, 2016.
- [119] H. Huijgens, M. Bruntink, A. van Deursen, T. van der Storm, and F. Vogelezang, "An exploratory study on functional size measurement based on code," *ICSSP '16 Proceedings of the International Conference on Software and Systems Process*, pp.56-65, 2016.
- [120] M. Ochodek, "Functional size approximation based on use-case names," *Information and Software Technology*, vol.80, pp.73-88, 2016.
- [121] S. Abrahao, L. De Marco, F. Ferrucci, J. Gomez, C. Gravino, and F. Sarro, "Definition and evaluation of a COSMIC measurement procedure for sizing Web applications in a model-driven development environment," *Information and Software Technology*, vol.104, pp.144-161, 2018.
- [122] A.Z. Abualkishik and L. Lavazza, "IFPUG Function Points to COSMIC Function Points convertibility: A fine-grained statistical approach," *Information and Software Technology*, vol.97, pp.179-191, 2018.

(2019年3月25日受付, 9月6日再受付,
11月20日早期公開)



山田 涼太

平成29年大阪大学基礎工学部情報科学科卒業。平成31年同大学大学院博士前期課程修了。ファンクションポイントや見積りについての研究に従事。



楠本 真二 (正員: シニア会員)

昭和63年大阪大学基礎工学部情報科学科卒業。平成3年同大学大学院博士課程中退。同年同大学基礎工学部助手。平成8年同講師。平成11年同助教授。平成14年同大学大学院情報科学研究科助教授。平成17年同教授。博士(工学)。ソフトウェアの生産性や品質の定量的評価, プロジェクト管理に関する研究に従事。IPSJ, JSSST, IEEE, JFPUG, PM学会, SEA, 各会員。



肥後 芳樹 (正員)

平成14年大阪大学基礎工学部情報科学科中退。平成18年同大学大学院博士後期課程修了。平成19年同大学院情報科学研究科コンピュータサイエンス専攻助教。博士(情報科学)。ソースコード分析, 特にコードクローン分析やリファクタリング支援に関する研究に従事。情報処理学会, 日本ソフトウェア科学会, IEEE 各会員。



裕本 真佑 (正員)

平成18年京都産業大学理学部卒業。平成22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。同年神戸大学大学院システム情報学研究科特命助教。平成28年大阪大学大学院情報科学研究科助教。博士(工学)。エンピリカルソフトウェア工学, 特にソフトウェアメトリクスの研究に従事。情報処理学会, IEEE, ACM 各会員。



倉重 誠

平成4年東京工業大学理学部情報科学科卒業。同年日立製作所に入社。平成30年株式会社SOFTEST設立。ソフトウェア開発・見積, プロジェクトマネジメントの各支援技術の開発, 普及推進に従事。FP技術者認定国際資格CFPS (Certified Function Point Specialist)の上級資格CFPS Fellowを国内・アジア初取得。日本ファンクションポイントユーザ会事務局長。プロジェクトマネジメント学会員。