

修士学位論文

題目

ファンクションポイントに関する研究論文の系統的レビュー

指導教員

楠本 真二 教授

報告者

山田 涼太

平成 31 年 2 月 6 日

大阪大学 大学院情報科学研究科

コンピュータサイエンス専攻

平成 31 年度 修士学位論文

ファンクションポイントに関する研究論文の系統的レビュー

山田 涼太

内容梗概

ソフトウェア開発の見積り元の元となる規模を評価する尺度の一つとして、ファンクションポイント (FP) が存在する。FP は 1979 年に A.J.Albrecht が考案した手法であり、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、様々な利点から現在も尚重要視されている。その一方で、FP に対して幾つかの問題点が指摘されており、これら問題に対しこの 40 年間で様々な研究が行われてきた。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この分野においてどのような既存研究が存在するのかを把握する事が難しい状況となっている。また、FP は見積もりにおいて現在も用いられる技術であるため、研究成果が開発現場において適用可能かどうかを知る事は非常に重要である。加えて、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はずれがあるという指摘があったため、FP という分野において研究と現場のニーズにギャップが存在していないかを確認する事は重要である。

本研究では、FP 研究における現在の進捗を収集し現在の課題と今後の研究の方向性を明らかにするため、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（以下、系統的レビューと呼ぶ）を実施した。系統的レビューを実施した結果、対象となった 95 本の論文において、6 つの研究トピックが存在する事や各トピックの研究内容とその成果、および研究動向が得られた。大半の研究において研究成果が適用可能となるプロジェクトのコンテキストについての記載が行われておらず、開発現場への適用を考える上での判断材料が不足している事が分かった。また、「発注側において FP が普及していない」や「FP についてのよい教育手法がない」といった現場のニーズがあるにも関わらず研究が行われていない分野があり、そのため開発現場と研究においてギャップがある事が分かった。

主な用語

ファンクションポイント法

系統的文献レビュー (Systematic Literature Review)

ソフトウェア工学

ソフトウェア開発見積もり

IFPUG 法

目次

1	まえがき	1
2	準備	3
2.1	ファンクションポイント法	3
2.2	Systematic Literature Review	5
2.3	Zelkowitz の評価モデル	6
3	実施した系統的レビュー	10
3.1	Research question の設定	10
3.1.1	RQ1	11
3.1.2	RQ2	11
3.1.3	RQ3	11
3.1.4	RQ4	11
3.2	研究論文の収集	11
3.2.1	検索する期間	12
3.2.2	検索対象データベース	12
3.2.3	検索に用いるキーワード	12
3.3	研究論文の選別	12
3.4	情報の抽出	13
3.5	結論の導出	14
4	レビュー結果	15
4.1	調査対象論文の外観	15
4.2	RQ1 への回答	15
4.2.1	計測補助	16
4.2.2	計測された FP の活用	17
4.2.3	計測ルールの変更	17
4.2.4	計測が難しいとされている対象への適用	18
4.2.5	その他	18
4.3	RQ2 への回答	19
4.4	RQ3 への回答	20
4.4.1	Observational Methods	20
4.4.2	Historical Methods	23

4.4.3	Controlled Methods	25
4.5	RQ4 への回答	26
5	考察	28
5.1	RQ1	28
5.1.1	FP の利点や欠点の評価	28
5.1.2	計測補助	28
5.1.3	計測された FP の活用	29
5.1.4	計測ルールの変更	30
5.1.5	計測が難しいとされている対象への適用	31
5.1.6	その他	32
5.2	RQ2	32
5.2.1	「評価」トピックの研究動向	33
5.2.2	「計測補助」トピックの研究動向	33
5.2.3	「活用」トピックの研究動向	34
5.2.4	「ルール変更」トピックの研究動向	35
5.2.5	「適用」トピックの研究動向	35
5.2.6	「その他トピック」の研究動向	36
5.3	RQ3	37
5.3.1	研究が適用可能なコンテキスト情報	37
5.3.2	検証に用いるプロジェクト	38
5.4	RQ4	39
5.4.1	汎用的な自動計測ツールがない	39
5.4.2	保守案件への適用が難しい	39
5.4.3	生産性評価モデルが確立されていない	39
5.4.4	発注側企業に FP が普及していない	40
5.4.5	FP で計測できない要素への対応	40
5.4.6	FP のよい教育手法がない	41
6	あとがき	42
	謝辞	43
	参考文献	44
	付録	55

表目次

1	FP 法に関する国際規格	3
2	ILF と EIF の複雑度	5
3	EI の複雑度	5
4	FP の算出のための計算表	6
5	各トピックに含まれる論文と本数	19
6	Zelkowitz の評価モデル	22
7	プロジェクトのコンテキスト	23
8	Observational Methods における検証データの出典	23
9	検証に用いた企業プロジェクトの数	24
10	6 件以上のプロジェクトを検証に用いた論文における提供元の企業数	24
11	Historical Methods におけるデータの出典	25
12	Historical Methods におけるデータの本数	25
13	Study of Legacy におけるデータの出典	26
14	Study of Legacy におけるプロジェクト数	26
15	開発現場のニーズと対応する論文の本数	27
16	「評価」トピックの研究内訳	29
17	「計測補助」トピックの研究内訳	30
18	「活用」トピックの研究内訳	30
19	「ルール変更」トピックの研究内訳	31
20	「適用」トピックの研究内訳	32
21	「その他」トピックの研究内訳	32
22	直近 15 年における各トピックの論文本数	37
23	直近 15 年における論文本数の多い小トピック (上位 3 つ)	38

目次

1	基本機能要素の例	4
2	論文収集と選別の結果	13
3	出版された年代ごとの調査対象論文数	20
4	出版された年代ごとの各トピックの割合	21
5	出版された年代ごとの「評価」トピックの論文数	33
6	出版された年代ごとの「計測補助」トピックの論文数	34
7	出版された年代ごとの「活用」トピックの論文数	35
8	出版された年代ごとの「ルール変更」トピックの論文数	36
9	出版された年代ごとの「適用」トピックの論文数	36
10	出版された年代ごとの「その他」トピックの論文数	37

1 まえがき

ソフトウェア開発の見積りの元となる規模を評価する尺度の一つとして、ファンクションポイント (FP) が存在する。FP は 1979 年に A.J.Albrecht が考案した手法であり、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る。FP には「ユーザ視点で機能を計測するため、ユーザとの合意が得られやすい」「ソフトウェアの機能を元に計測するため、実装に用いる開発言語に依存しない」等の利点から現在もなお重要視されており、政府情報システムの開発においても、要求内容に設計または開発に関する工程が含まれる場合には、原則 FP の見積り及びその根拠を示すことが必須となりつつある [4]。

その一方で、FP には「主観的判断が含まれるため計測結果に対する信頼性に不安が残る」「計測の手間が大きい」などの様々な問題が指摘されており、これらの問題に対しこの 40 年間で「信頼性の評価」や「自動計測手法の提案」、「計測された開発規模と工数の関係性の導出」など様々な研究が行われてきた。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。FP の計測上における問題点とその改善案についての二次研究は行われているが [116]、FP 研究全体を対象とした二次研究は未だ行われていない。そのため、新規の研究者や開発者がこの分野においてどのような既存研究が存在するのかを把握する事が難しい状況となっている。また、FP は見積りにおいて現在も用いられる技術であるため、研究成果が開発現場において適用可能かどうかを知る事は非常に重要である。加えて、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ずれがあるという指摘があったため、FP という分野において研究と現場のニーズにギャップが存在していないかを確認する事は重要である。

以上の背景から、本研究では、FP 研究における現在の進捗を収集し現在の課題と今後の研究の方向性を明らかにするため、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（以下、系統的レビューと呼ぶ）を実施した。系統的レビューとは、網羅的で再現性のある文献調査手法である [9, 10, 11]。系統的レビューは主に、定式化した課題について既存研究を元に回答を行うために実施される。本研究では、レビューを行うに当たって明らかにする Research Question を 4 つ設定した。1 つ目 (RQ1) は、FP 研究にはどのようなトピックが存在するか、である。2 つ目 (RQ2) は、FP 研究の推移はどのようなものか、である。3 つ目 (RQ3) は、研究成果は現場での適用が可能か、である。4 つ目 (RQ4) は、現場のニーズと研究にギャップは存在するか、である。主要国際会議の論文集と主要論文誌に採録された論文 95 本を対象に系統的レビューを実施した。RQ1 への回答として、「計測された FP の活用」「FP の計測ルールの変更」「FP の計測補助」「特定の開発手法やシステムへの FP の適用」「FP 法の評価」「その他」の 6 つのトピックの研究が存在する事が分かった。RQ2 への回答として、「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」についての研究から「計測補助」についての研究へと研究内容が推移した事が分かった。RQ3 への回答として、大半の研究において研究成果が適用可能となるプロジェクトのコンテキストについての記載が行われておらず、開発現場への適用を考える上での判断材料が不足している事が分かった。RQ4 への回答として、開発現場のニーズと研究にギャップが存在する事が分かった。「発注側においてFPが普及していない」や「FPについてのよい教育手法がない」といった現場のニーズがあるにも関わらず研究が行われていない分野があり、そのため開発現場と研究においてギャップがある事が分かった。

以降、2章では本研究において前提知識となる用語の解説を行う。3章では系統的文献レビューの手順について詳細に説明を行う。4章では系統的文献レビューを実施して得られた情報からRQへの回答を行う。5章では各RQの回答に対する考察を行う。最後に6章で本研究のまとめと今後の課題について述べる。

表 1: FP 法に関する国際規格

ISO IEC 国際規格	手法	発行年
20926	IFPUG 法	2009
20968	Mk II 法	2002
24570	NESMA 法	2005
19761	COSMIC	2011
29881	FISMA 法	2010

2 準備

本章では研究の背景となる諸用語や関連研究について簡単に述べる。

2.1 ファンクションポイント法

ソフトウェアの規模を見積もる手法の 1 つにファンクションポイント (FP) 法 [3] がある。FP 法は、ユーザから見た機能の量を計測する手法で、A.J.Albrecht によって 1979 年に提案された。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。以降、IFPUG 法 [5]、MarkII 法 [6]、COSMIC 法 [7] 等、様々な FP 法が提案されてきている。FP 法に関する国際規格を表 4 に示す。FP 法の中で、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 法において、ソフトウェアの持つ機能が分類される 5 種類の基本機能要素を以下に示す [8]。

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持ったデータの集合

外部インターフェイスファイル (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、出力データに導出データを含むもの

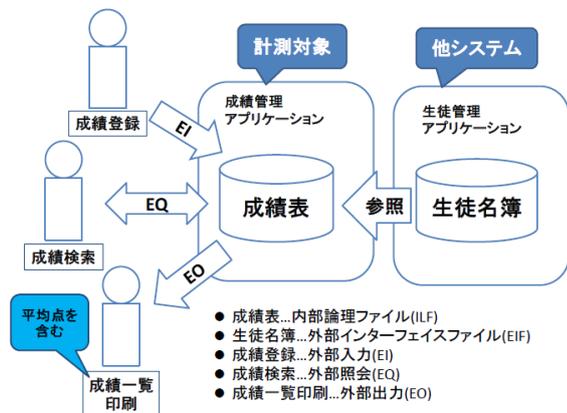


図 1: 基本機能要素の例

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、出力データに導出データを含まないものであり、処理が ILF を更新しないもの

例として、図 1 に学生の成績管理アプリケーションとそこに含まれる基本機能要素を示す。計測対象である成績管理アプリケーション内でデータが更新される成績表は ILF、成績管理アプリケーション内にはなく、成績管理アプリケーションから参照される生徒名簿は EIF となる。また、データ入力によって成績の登録を行う成績登録処理は EI、成績表から平均点などの派生データを含まないデータを出力させる成績検索は EQ、成績表から平均点などの派生データを含むデータを出力させる成績一覧印刷は EO となる。上記の 5 種のうち、ILF と EIF はデータファンクション (以降、DF)、EI と EQ と EO はトランザクションファクション (以降、TF) と呼ばれ、それぞれ計測に必要な要素が異なる [8]。

- DF
 - 種別…ILF, EIF
 - DET(Data Element Type)…アプリケーション上で参照される項目の数
 - RET(Record Element Type)…データ群の属性内のサブグループの数
- TF
 - 種別…EI, EQ, EO
 - DET…重複を除いた入出力項目の数に、該当処理において下記のものがあれば 1 ずつ増やしたもの

表 2: ILF と EIF の複雑度

RET \ DET	1-19	20-50	> 50
1	低	低	中
2-5	低	中	高
> 5	中	高	高

表 3: EI の複雑度

FTR \ DET	1-4	5-15	> 15
0-1	低	低	中
2	低	中	高
> 2	中	高	高

* きっかけ (処理が行われるトリガー、ボタンなど)

* メッセージ (エラーメッセージなど)

- FTR (File Type Reference) … TF に関わる DF の数

上述した要素を用いて各機能の複雑度を計測し、その上で FP の値を求める。

- DF の複雑度

表 2 を参照し DF の複雑度を決定する。例えば RET が 1、DET が 26 の ILF があった場合、表 2 より複雑度は低となる。

- TF の複雑度

EI、EO、EQ のそれぞれに対して用意されている対応表を用いて、TF の複雑度を決定する。例えば、EI は、表 3 を参照し複雑度を求める。FTR が 2、DET が 14 の EI の場合、複雑度は中となる。

- 複雑度からの FP 計測

抽出された基本要素の複雑さ毎の個数を表 4 に代入し、それぞれに設定されている重み数を掛けて足し合わせた値が FP となる。前述の例であれば、複雑度「低」の ILF (RET : 1, DET : 26) であれば重みは 10 となる。また、複雑度「中」の EI (FTR : 2, DET : 14) であれば重みは 4 となる。

2.2 Systematic Literature Review

系統的文献レビュー (Systematic Literature Review) とは、網羅的で再現性のある文献調査手法である [9, 10, 11]。SLR では明確な手順を踏んで調査を行うため、レビューそのものの再現性が高い。そのため、手順の設定が明確に行われていればレビューの実行者に関わらず同様の結果を得る事ができ、レビュー結果の信頼性は高いとされている。系統的文献レ

表 4: FP の算出のための計算表

	低	中	高	合計
ILF	$\square \times 7 = \square$	$\square \times 10 = \square$	$\square \times 15 = \square$	
EIF	$\square \times 5 = \square$	$\square \times 7 = \square$	$\square \times 10 = \square$	
EI	$\square \times 3 = \square$	$\square \times 4 = \square$	$\square \times 6 = \square$	
EO	$\square \times 4 = \square$	$\square \times 5 = \square$	$\square \times 7 = \square$	
EQ	$\square \times 3 = \square$	$\square \times 4 = \square$	$\square \times 6 = \square$	
			FP	

ビューは医療の研究分野において、情報収集や調査を行う際によく用いられる手法であり、ソフトウェア工学研究分野においては、主に提案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。ソフトウェア工学研究分野における系統的レビューの手法は Kitchenham により提案されており [9]、この手法に従った系統的レビューが盛んに実施されている [12]。本研究で実施する系統的レビューも概ね Kitchenham らが提案した手法に従っている。見積もりの研究分野においては、機械学習を利用した見積もり手法に関する系統的レビュー [13] や類推見積もり手法に関する系統的レビュー [14]、FP の計測上の問題点と改善案に関する系統的レビュー [116] が行われている。系統的レビューの具体的な手順については、3 章で詳細な説明を行う。

2.3 Zerkowitz の評価モデル

Zerkowitz 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検証方法を 12 種類に分類したものである [25]。分類としては、大分類として 3 種類、小分類として 12 種類の分類が存在する。下に、Zerkowitz の評価モデルの詳細を示す。

- 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・検証する。

- Project monitoring

単一のプロジェクトに研究成果を適用し、その結果を分析する。

- * 利点…他検証方法と比べて実施のためのコストが低い
- * 欠点…従来手法等との比較は行わないため、検証によって示される研究の妥当性が低い

– 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。

- * 利点…単一のプロジェクトでのみ検証を行うため、実施のためのコストが低い
- * 欠点…後述の Assertion や Field study と比べると検証で示される研究の妥当性が低い

– 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、従来手法などの結果と比較する事で、一定の規模以内のプロジェクトに対し研究成果が有効であるかどうかを評価する。

- * 利点…一定規模以下とはいえ複数のプロジェクトで検証を行うため、単一のプロジェクトでのみ有効性を評価する Case study と比べると検証で示される研究の妥当性が高い
- * 欠点…Field study と比べると検証によって示される妥当性は低い

– 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、研究成果が普遍的に有効であるかを評価する。

- * 利点…規模に制限はなく対象を幅広く取るため、一定規模以下のプロジェクトで検証を行う Assertion と比べると検証によって示される研究の妥当性は高い
- * 欠点…複数プロジェクトで検証を行うため、実施のためのコストが高い

● 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。Observational Methods や Controlled Methods と異なり、プロジェクトを用いて研究成果を検証するのではなく、プロジェクト等の分析から研究成果を得る研究がこの分類に含まれる。

– Literature search

過去の文献を元に知見を得る。SLR 等が当てはまる。

- * 利点…実施のためのコストが低い
- * 欠点…文献の選択によってバイアスが生じる場合がある

– 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。

- * 利点…実施のためのコストが低い
- * 欠点…用いるデータによってバイアスが生じる場合がある
- 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。
 - * 利点…実施のためのコストが低い
 - * 欠点…定量的なデータではない
- 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。Study of Legacy が工数や開発規模のような数値的なデータを分析するのに対し、Static analysis は仕様書やソースコードのような成果物を分析する。
 - * 利点…ツールによって自動分析が可能な場合は実施のためのコストが低い
 - * 欠点…開発手法を評価する事はできない
- 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・検証を行う。
 - 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。Field study と比べ、検証において比較したい内容が厳密に定められており、それ以外の内容で差が生じないよう検証に用いられるプロジェクトや適用方法等も厳密に管理される。
 - * 利点…Field study と比べて検証によって示される研究の妥当性が高い
 - * 欠点…実施のためのコストが高い
 - Synthetic Environment Experiments
 - あるプロジェクトに対し、検証グループの半分は提案手法を、もう半分には従来手法を用いさせ、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。
 - * 利点…複数人の検証によって人的要因が消去されるため検証によって示される研究の妥当性が高い
 - * 欠点…実施のためのコストが高い
 - 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。

- * 利点…ツールによって自動分析が可能な場合は実施のためのコストが低い
- * 欠点…開発手法を評価する事はできない

– **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。

- * 利点…ツールによって自動分析が可能な場合は実施のためのコストが低い
- * 欠点…データが現実にそぐわない場合がある

本研究ではこの評価モデルをFPについての研究成果が評価されているか、ひいては開発現場での適用が可能であると言えるほど妥当性を獲得しているかを検証するために使用する。

3 実施した系統的レビュー

本章では，系統的レビューの手順について詳細に説明を行う。

本研究では，Kitchenham らが提案した手法に従い系統的レビューを行う。レビュー手順は以下に示す。

1. Research Question(RQ) の設定

レビューで明らかにしたい問いを設定する

2. 研究論文の収取

調査対象となる論文を収集する予め検索対象，キーワード，年代を設定する

3. 研究論文の選別

調査対象として適していない論文を除外する

4. 情報の抽出

RQ への回答に必要な情報を研究論文から抽出する

5. 結論の導出

抽出した情報を元に RQ への回答を行う

本研究で行う系統的レビューは，Research question の設定，研究論文の収集，研究論文の選別，情報の抽出，結論の導出の 5 つの手順から成る。以降，各手順について説明を行う。

3.1 Research question の設定

この手順では，系統的文献レビューを行うにあたり明らかにしたい問いを Research question (RQ) として設定する。本研究では，FP に対してどのような研究が行われているのかと今後どのような研究を行えばよいのかを明らかにするために，以下の 4 つの Research question を設定した。

- RQ1:FP 研究にはどのようなトピックが存在するか
- RQ2:FP 研究の推移はどのようなものか
- RQ3:研究成果は現場での適用が可能か
- RQ4:開発現場のニーズと研究にギャップは存在するか

3.1.1 節で RQ1 について，3.1.2 節で RQ2 について，3.1.3 節で RQ3 について，3.1.4 節で RQ4 についての説明をそれぞれ行う。

3.1.1 RQ1

FPについての研究はFPが提案された1979年から現代に至るまで行われている。40年間で行われた研究は膨大である一方で、FP全般についての研究を整理した研究は存在していないため、FPという分野でどのような研究がこれまで行われてきたかを網羅的に把握する事は困難である。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調査する。

3.1.2 RQ2

上述の研究内容と同様に、40年間で行われた研究の内容がどのように移り変わっていったかを把握する事は現状困難である。既存研究の研究動向を明らかにするため、FPが提案された1979年から現代に至るまで、FPという分野における研究内容がどのように推移したかを調査する。

3.1.3 RQ3

FPはソフトウェアの開発見積もりに利用される尺度であるため、ソフトウェア開発の技術の中でも特に開発現場で活用される技術であると言える。そのため、研究成果が開発現場で適用可能かどうかは非常に重要な事柄である。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検証されているかを調査する。

3.1.4 RQ4

ソフトウェア工学の分野において、開発現場のニーズと実際に行われている研究にはギャップが存在するという研究が発表されている [27]。この論文では、ソフトウェアの開発者が「正確な見積もり手法がない事」などのソフトウェアの生産性に関わるニーズを抱いているのに対し、研究としては生産性よりもテストやデバッグなどのソフトウェアの品質についての研究が多く行われている現状があり、開発現場のニーズと研究に乖離が生じていると結論付けた。開発現場のニーズと研究に乖離が存在した場合、研究が行われていないニーズについては今後研究が行われる必要がある。今後研究が必要なニーズの有無を確認するため、FPという分野において開発現場のニーズと研究にギャップが生じているかどうかを調査する。

3.2 研究論文の収集

この手順では、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、および検索に用いるキーワードを

設定する。

3.2.1 検索する期間

過去 40 年間 (1979 年から 2018 年), FP が提案されてから今日までを対象とした。

3.2.2 検索対象データベース

検索対象としたデータベースは以下の 5 つのデータベースである。

- IEEE xplora [18]
- ACM Digital library [19]
- Science Direct [20]
- Scopus [21]
- Google Scholar [22]

見積り研究分野における研究論文が多く含まれているため, 上記データベースを検索対象として選択した。

3.2.3 検索に用いるキーワード

本研究では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。

Google Scholar 以外の 4 データベースに対しては, タイトル, 内容梗概, キーワードのいずれかに上記のキーワードが含まれる論文を出力するよう検索オプションを設定したが, Google Scholar に対してはタイトルにのみ上記キーワードが含まれる論文を出力するよう検索オプションを設定した。これは, 内容梗概とキーワードに上記キーワードが含まれる論文を出力するよう設定すると, 膨大な量の無関係な論文が出力されるためである [13]。

3.3 研究論文の選別

この手順では, 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 (医療分野の論文など) を候補から取り除く

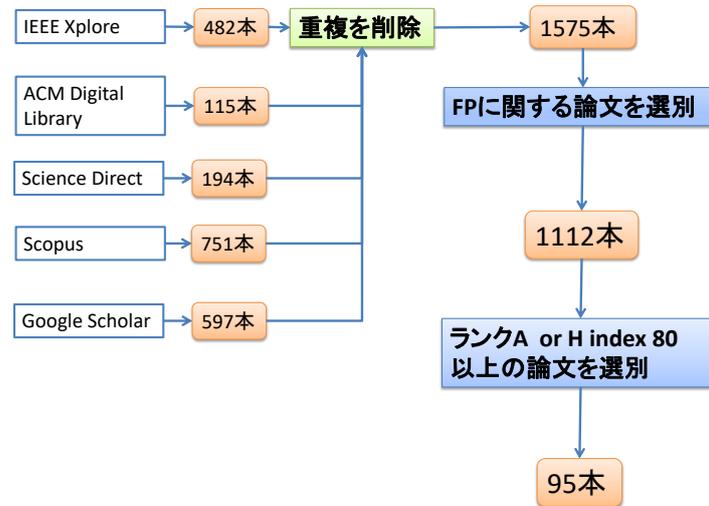


図 2: 論文収集と選別の結果

3. 各論文の出典を確認し，CORE RankingsPortal [15] 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および H index 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

Kitchenham らの提案手法 [9] に従った系統的レビューでは，上記手順 3 において候補となっている論文の本文を確認し，論文の質の評価を行っている系統的レビューが主だが，質の評価を上記手順 3 のように，論文の出典で質の評価を行っている系統的レビューもいくつか存在する [17]．本研究ではそのような論文に倣い，質の評価を論文の出典で行った．上記選別手順を実行した結果を図 2 に示す．収集した論文候補から重複を取り除いた所，1575 本の論文が得られた．この候補から FP に関連しない論文を取り除くと 1112 本が候補として残った．1112 本の候補から，CORE RankingsPortal [15] 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および H index [16] 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いた結果，95 本が調査対象の論文として選択された．調査対象として選択された論文のタイトルと著者名，出典名，出版年を「付録 A」に示す．

3.4 情報の抽出

この手順では，調査対象とした研究論文の本文を確認し，論文に関する基本的な情報と Research Question への回答に必要な情報を抽出する．論文に関する基本的な情報とは，タイ

トル、著者名、出典名、出版年である。抽出する情報を以下に示す。

- タイトル
- 著者名
- 出典名
- 出版年
RQ2 において研究の推移を調査するために必要。
- 研究の背景
RQ2 において研究の推移の原因を分析するために必要。
- 研究の目的
RQ1 において研究をトピックに分類するために必要。
- 手法
RQ1 において研究内容を整理・分析するために必要。
RQ4 において研究と開発現場のニーズを照らし合わせるために必要。
- 結果
RQ3 において研究がどのように検証されているかを分析するために必要。
- 結論
RQ1 において研究成果を整理・分析するために必要。

3.5 結論の導出

この手順では、抽出した情報を整理・考察した上で Research Question への回答を行う。RQ1 への回答のためには、既存研究のトピックごとの分類のために、対象となった既存研究の内容を網羅的に把握する必要がある。RQ2 への回答のためには、FP 分野における研究の推移を調査するために、年代別の研究が盛んなトピックの把握が必要である。RQ3 への回答のためには、研究成果が開発現場で適用可能かを調査するために、研究成果が十分に検証されているかを確認する必要がある。RQ4 への回答のためには、開発現場のニーズと研究との間にギャップが存在するかを調査するために、開発現場のニーズと既存研究を照らし合わせる必要がある。各集計結果と Research Question への回答については 4 章、その結果を受けた考察については 5 章で説明する。

4 レビュー結果

4.1 調査対象論文の外観

3章で説明した通り、論文の収集と選別を行った結果95本の論文が調査対象として選択された。各調査対象論文のタイトル、著者名、出典名、出版年を「付録A 調査対象論文」に示す。以降の小節で各RQへの回答を行う。

4.2 RQ1への回答

既存研究を分析し、95本の論文を6つのトピックに分類した。分類された論文とその本数を表5に示す。また、今回対象とした95本の論文の中には「FPの利点や欠点の評価」と「その他」の2つのトピックに関する研究が1本存在したが[122]、研究の主目的が「IFPUG法とCOSMIC法の比較」であったため、「評価」トピックの研究として分類した。

- 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。

- 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。計測自体は可能であるシステムや開発手法に対して補助を行う研究がこのトピックに含まれる。そもそも計測が難しいとされている開発手法などからの計測は後述の「適用」トピックに含まれる。

- 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の結果を見積も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。

- 計測ルールの変更

FPの計測ルールの校正も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、既存の計測ルールを改善するために行われている研究は「変更」がこのトピックに含まれる。

- 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。従来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上述の「評価」トピックに含まれ、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「計測補助」トピックに含まれる。

- 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. 上記の 5 つのトピックに当てはまらない研究がこのトピックに含まれる。

「他の開発規模尺度と比べてどういう点で優れているのか」について研究している論文は「評価」, 「どうすれば計測の手間が減らせるか」について研究している論文は「計測補助」, 「計測によって得られた FP をどのように工数に変換するか」を研究している論文は「活用」トピックにそれぞれ分類される。また, 「どのように重みづけを較正すればより正確な FP が得られるか」を研究している論文は「変更」, 「どうすれば FP を得られるようになるか」を研究している論文は「適用」としてそれぞれ分類される。

各トピックにおいて引用数が多い上位 2 本の論文の概要を示す。

4.2.1 計測補助

- Diab らの研究 [77]: モデル駆動型開発ツールにおける COSMICFP の自動計測

電気通信や航空電子工学などで使用されていたモデル駆動型開発ツールである Rational Rose RealTime(RRRT) における COSMIC 法の自動計測が研究された。計測のために RRRT のモデルにおける要素と COSMIC 法の計測要素が紐づけられ、自動計測手法としてツールに実装された。自動計測結果が専門家との計測結果と比較され誤差が生じたものの、仕様に対する解釈の違いのために生じた誤差であり計測結果としては許容範囲内であった。結論として、提案された自動計測手法の計測結果が許容範囲内である事、および RRRT において管理されるシステムの要素は非常に詳細であり人の手による計測では情報が見落とされる可能性があるため自動計測は RRRT に対し非常に有用であるとされた。

- Zivkovic らの研究 [82]:UML からの FP 自動計測

計測の手間の削減やオブジェクト指向への FP 法の適用のため、UML からの FP 計測手法が提案された。開発が進むにつれて見積もりに利用可能な情報が増えていく事を考慮し、計測手法として「ユースケース図からの計測」「アクティビティ図とシーケンス図からの計測」「クラス図からの計測」という情報量の異なる三段階の計測が提案され、その内「アクティビティ図とシーケンス図からの計測」「クラス図からの計測」についてはツールによる計測の自動化が実現された。計測結果を比較した結果、情報が増加するにつれて計測精度は向上する一方で「ユースケース図からの計測」と「ア

クティビティ図とシーケンス図からの計測」の計測精度にはほとんど差が生じなかったため、提案手法を利用する際は「ユースケース図からの計測」と「クラス図からの計測」のみでよく、「アクティビティ図とシーケンス図からの計測」は行う意味がないと結論付けられた。

4.2.2 計測された FP の活用

- Matson らの研究 [49]:FP から工数を見積もる統計回帰モデル

FP から工数を見積もるための統計回帰モデルの制作を行い、その精度を評価した。104 のプロジェクトのデータをもとに線形回帰モデルと重回帰モデルが開発され、それぞれの見積もり精度として実際の工数との誤差が調査された。結果、25%以内の誤差に収まった値はそれぞれ 64%と 68%となり、目標としていた 75%にはわずかに届かなかった。

- Fnnie らの研究 [56]:FP を用いた工数見積もり手法の比較

FP を使用した工数見積もり手法として存在する回帰分析、ニューラルネットワーク、類推法の比較によりどの見積もり手法が優れているかが評価された。299 件のプロジェクトを用いた比較の結果、ニューラルネットワークと類推法において、実工数との MARE(平均絶対相対誤差) がそれぞれ 0.352 と 0.362 であり高い見積もり精度が示された一方で、回帰分析による工数見積もりは MARE が 0.623 と他 2 手法と比べて低い計測精度であった。

4.2.3 計測ルールの変更

- Finnie らの研究 [56]:より複雑度と調整係数が少ない FP 計測手法の提案

「低・中・高」の 3 種類である複雑度を「中」に統一し、本来 14 種類存在する調整係数を「論理的な複雑さ」と「データの複雑さ」のみにする事により、計測に必要な情報量を削減した新計測手法「SPQWZO FP」の提案およびその計測精度の評価が行われた。検証に用いた 64 件のプロジェクトの内、提案手法は従来の FP に対し 38 件において誤差 20%以内の FP を計測し、40%以上の誤差が生じたのは 4 件のみであった。そのため、提案手法は有用であると結論付けられた。

- Wei らの研究 [90]:統計解析やニューラルネットワークによる重みづけの再定義

重みづけの是正のため、統計解析とニューラルネットワーク、そしてファジィ論理を用いた重みづけの再定義が行われた。再定義の結果全ての重みの値が小さくなり、これは FP が提案されて以降の 25 年間で生産性が高まりタスクに対する工数が減ったた

めだと考察された。また、新たに定義された重みづけを用いて FP が計測され、従来の FP との工数見積精度が比較された。MMRE において平均 22% の改善が見られたため、重みづけの較正による見積もり精度の向上が示された。

4.2.4 計測が難しいとされている対象への適用

- Reifer らの研究 [38]:科学的及びリアルタイムシステムへの FP の適用

科学的及びリアルタイムシステムには複雑な内部処理が存在するため FP の適用が難しいとされている。既存研究として、FP に加えて「技術的要因」や「言語拡張係数」などの非機能要件を用いる事で科学的及びリアルタイムシステムの開発規模 (SLOC) を見積もる ASSET という手法が存在し、この研究においては ASSET を自動的に計測できる ASSET-R が開発され、現場へ導入する事でその有用性の評価が行われた。3 年間の実験期間において、ASSET-R は 34 件のプロジェクトに適用され、誤差 20% 以内の行数見積もりが可能である事が示された。

- Hericko らの研究 [93]:反復型開発への FP の適用

反復型開発では反復のたびに仕様が追加され、プロジェクトの最後にプロジェクト全体の仕様が完成するため、プロジェクト全体の仕様を元に計測を行う FP 法は適用が難しいとされている。反復型開発への FP の適用のため、各反復における最初と最後の仕様に対し計測を行い、その誤差を次の反復での見積もりに利用する手法が提案された。提案手法は 3 つのプロジェクトによって検証され、3 つの内 2 つにプロジェクトにおいて 2 度目の反復における工数見積もりにおける誤差が 10% 以下となり、その有用性が示された。

4.2.5 その他

- Cuadrado-Gallego[92]:

IFPUG 法で得られた過去データを COSMIC 法にも適用可能とするため、IFPUG 法の FP を COSMIC 法の FP に変換するための研究が行われた。手法としては IFPUG 法における各 TF の FTR 数を元に COSMIC 法における FP の下限と上限を導出する事で、IFPUG 法における FP を COSMIC 法の FP に変換する。33 のプロジェクトを用いて、導き出された上限・下限と実際の COSMIC 法の FP が比較された結果、それぞれの差は指数分布に従ったため、提案手法は適用可能であると結論付けられた。

- Cuadrado らの研究 [87]:

IFPUG 法と COSMIC 法の変換手法を研究するためには、双方の計測結果が必要で

表 5: 各トピックに含まれる論文と本数

トピック	含まれる論文	本数
FP の利点や欠点の評価	[32, 34, 36, 37, 39, 42, 43, 44, 45, 47, 48, 50, 51, 54, 57, 59, 66, 68, 70, 81, 85, 94, 96, 99, 105, 109, 112, 122]	28
計測補助	[58, 60, 69, 72, 73, 74, 77, 78, 82, 83, 84, 86, 88, 89, 97, 101, 102, 104, 106, 113, 115, 121, 123, 124, 125]	25
計測結果の活用	[33, 49, 52, 53, 55, 56, 61, 65, 91, 98, 103, 108, 110, 117]	14
計測ルールの変更	[35, 40, 41, 46, 63, 64, 67, 76, 79, 80, 90, 100, 111, 118]	14
計測が難しいとされている対象への適用	[38, 62, 71, 75, 93, 95, 114, 119, 120]	9
その他	[87, 92, 107, 116, 126]	5

ある。しかしながら既に普及している IFPUG 法と違い、手法として普及が進んでいない COSMIC 法の計測結果を収集するためには計測の専門家の協力が必要であり、人件費が高くてついでしまう。また、不適切な計測訓練を行った場合や異なる計測者による計測結果を計測者による誤差を考慮せずにまとめてしまった場合、計測データの品質が悪く、誤った変換手法が導出される可能性が高い。そのため、人件費の安い学生に計測を行わせる事でデータの収集コストを抑え、また授業への出席率や書面テストの正答率から評価の高い学生を集める事によってデータの品質を高める試みが行われた。收拾されたデータをもとに IFPUG 法と COSMIC 法の変換モデルを制作した結果、従来の線形モデルと同じく傾きが 1 に近い変換モデルが得られた。この結果から、信頼性の高いデータセットの取得のためには学生を計測者として安価に用いる手法が有用であると示された。

4.3 RQ2 への回答

この節では FP 研究における研究動向について分析を行った。分析のために、40 年間の研究をスタージェスの公式 [26] に則り 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 3 に、年代ごとの各トピックの研究の割合を図 4 に示す。

前半 20 年、特に 1989 年以降において「評価」トピックと「活用」トピックについての研究が盛んに行われている。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事で研究が落ち着いた

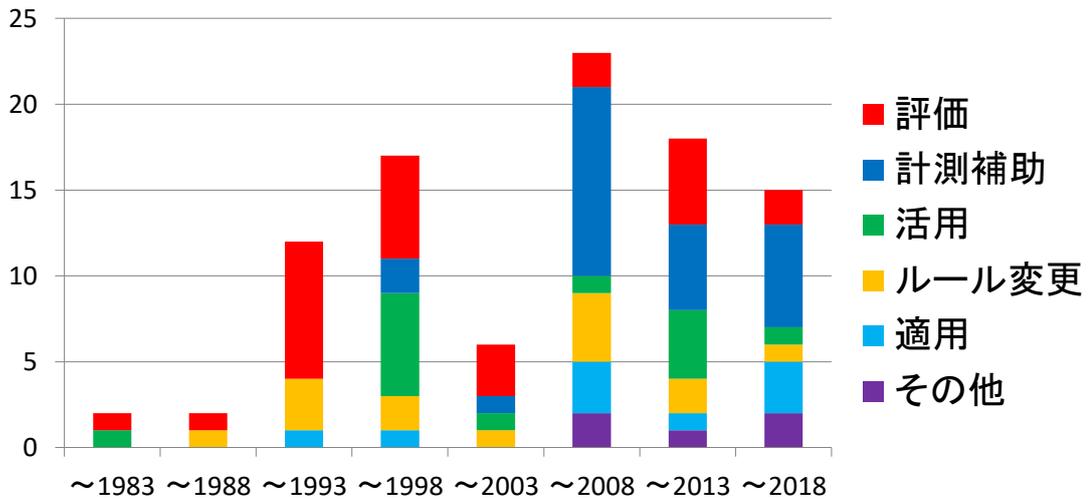


図 3: 出版された年代ごとの調査対象論文数

ためだと考えられる。また、この期間においては 2004 年以降に研究数を増やしており、特に「計測補助」についての研究が盛んである。「計測補助」トピックの研究が増加した理由は、2003 年に XML Metadata Interchange(XMI)[28] が登場し、Unified Modeling Language(UML) を電子的に扱う事が可能となった事から UML[29] からの自動計測が盛んに研究されるようになったためだと考えられる。各トピックにおける研究動向とその考察については 5.2 章で述べる。

4.4 RQ3 への回答

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、各研究を Zelkowitz の評価モデルを用いて分類した。そして、各分類ごとに研究がどのように評価されているかを分析した。Zelkowitz の評価モデルに則り研究を分類した結果を表 6 に示す。今回対象とした論文の中には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検証が行われていない論文が含まれていたため、「No experimentation」という分類を作成した。

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の回答を行う。

4.4.1 Observational Methods

Observational Methods に分類された研究では、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た。自動計測手法についての研究の場合は、従来の手法で計測された FP と比較する事で提案手法が従来手法の代替として実行できるかを評価し、計測ルールの変更の場合は提案手法によって得られた FP から見

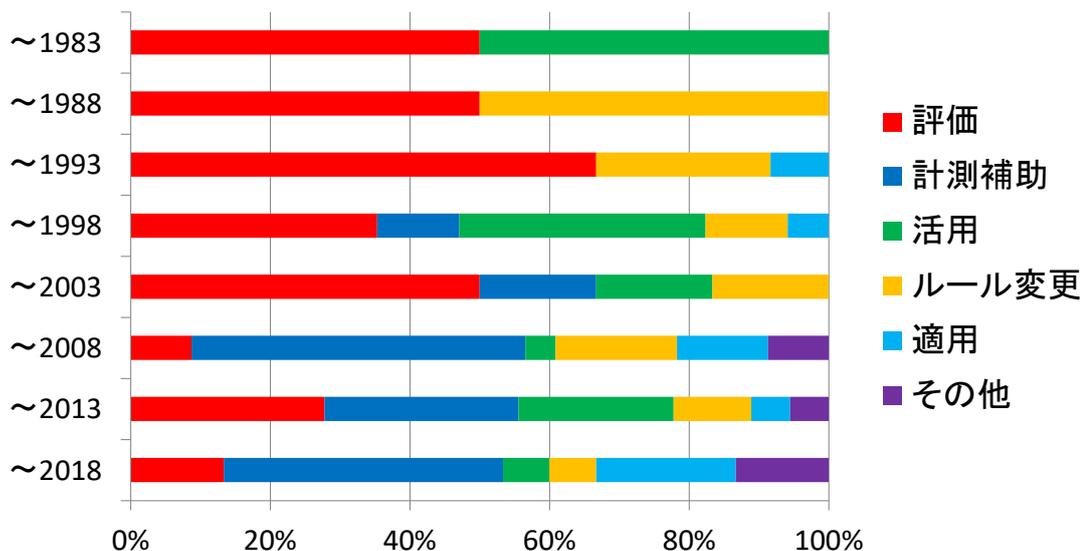


図 4: 出版された年代ごとの各トピックの割合

積もった工数と従来の手法で計測された FP を元に見積もった工数，そして実際にかかった工数を比較する事で提案手法が従来の手法より良い計測精度であるかどうかを評価していた。

研究成果の開発現場での適用可能性を考える上で，研究成果が適用可能な対象のコンテキストについての情報が重要である。多くの研究において，研究成果は一定のコンテキストを持つプロジェクトで適用が可能である場合が多い。ここで言うコンテキストとは，研究成果が適用可能なプロジェクトにおける開発言語や開発規模，業種等である。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持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。そのため，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た場合，それらの情報は開発現場に研究成果を導入する判断材料となる。逆に，コンテキストについての情報が記載されていない場合，研究成果を導入する判断材料が不足しているとして現場への適用が見送られる可能性が高いため，研究論文にはそれらのコンテキストが記載されている必要があると言える。そういった背景から，Observational Methods に分類された論文において，研究成果が適用可能なプロジェクトのコンテキストが十分記載されているかどうかを調査した。今回は IPA SEC が発行しているデータ白書 [31] の企業プロジェクトの収集において取得されるデータ項目をプロジェクトのコンテキストとして選択し，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 7 に示す。Observational Methods に分類された論文において，コンテキスト項目が全て記載されて

表 6: Zelkowitz の評価モデル

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

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すべ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言える。

また、研究成果の開発現場での適用可能性を考える上で「研究成果が十分に検証されているか」も重要な要素である。例えば研究成果が少数のプロジェクトでのみ検証されていた場合、その妥当性は低いものであると言うことができ、開発現場での適用可能性は低いと考えられる。そのため、本分類の研究において検証に用いたプロジェクト数などを調べた。各研究における検証に用いたデータの出典およびプロジェクト数をそれぞれ表 8 および表 9 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検証されているため、実際の開発現場に即した環境で検証は行われていると言える。一方で、企業のプロジェクトを用いて検証されている研究のうち、46 本中 17 本において検証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。これらの研究は研究の妥当性が十分検証されていないために開発現場への適用可能性は保証されていないと言える。また、検証のために用いられたプロジェクト数が 6 件以上であるものの、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 本において存在した。これらの研究成果はプロジェクトを提供した企業以外の他企業における妥当性が示されていないため、開発現場への適用可能性は保証されていないと言える。そのため、現状開発現場への適用可能性が保証されていると言えるのは、検証のために用いられたプロジェクト数が 6 件以上でなおかつプロジェクトを複数の企業から得ている研

表 7: プロジェクトのコンテキスト

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

表 8: Observational Methods における検証データの出典

データの出典	論文数
企業	46
大学	5
ISBSG	3

究 7 本のみであり, そのため Observational Methods に分類された研究全てが開発現場での適用可能性が十分であるとは言い難い.

4.4.2 Historical Methods

Historical Methods に分類された研究では, 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. そのため, 前述の Observational Methods や Controlled Methods のように研究成果の検証は行われておらず, 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およびプロジェクト数をそれぞれ表 11 および表 12 に示す.

小分類の違いが「検証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検証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 件は既存研究として紹介した 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

表 9: 検証に用いた企業プロジェクトの数

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

表 10: 6 件以上のプロジェクトを検証に用いた論文における提供元の企業数

企業数	論文数
1	21
2～7	7

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および件数に依存すると考えられる。これは、例えば大学の演習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、少数のプロジェクトから得られた知見には普遍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およびプロジェクト数をそれぞれ表 13 および表 14 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 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

表 11: Historical Methods におけるデータの出典

データの出典	論文数
企業	20
ISBSG	4
研究論文	3
大学	1

表 12: Historical Methods におけるデータの本数

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

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と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。

4.4.3 Controlled Methods

Controlled Methods に分類された研究では、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、その厳密な検証方法が挙げられる。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 は検証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検証している。しかしながら、複数の人々に検証させる事によって人的要因を無くし、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が

表 13: Study of Legacy におけるデータの出典

データの出典	論文数
企業	15
ISBSG	4
大学	1

表 14: Study of Legacy におけるプロジェクト数

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

ら検証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検証が困難であり、4件の研究においても検証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」を検証する側面が強く、後者は「提案手法は誰が行っても同じ結果が得られるか」を検証する側面が強い。そのため、研究によってどちらが検証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検証手法だと言える。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および提案手法の評価は難しいと考えられる。

4.5 RQ4 への回答

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)の有識者会議によって得られたFPに対する開発現場のニーズ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。得られたニーズと対応する論文の本数を表15に示す。6つのニーズの内「自動計測ツールがない」「保守案件への適用が難しい」「生産性評価モデルが確立されていない」「発

表 15: 開発現場のニーズと対応する論文の本数

開発現場のニーズ	対応する研究	論文数
自動計測ツールがない	[58, 60, 69, 72, 73, 77, 78, 82, 83, 89, 102, 104, 121, 123, 124]	15
保守案件への適用が難しい	[62, 119, 120]	3
生産性評価モデルが確立されていない	[33, 65]	2
FPで計測できない要素への対応	[110]	1
発注側企業にFPが普及していない	[117]	1
FPのよい教育手法がない		0

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5つのニーズにおいて対応する研究数が3本以下であり、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、開発現場と研究にはギャップが存在すると言える。各ニーズに対する考察は6.4節で述べる。

5 考察

この章では、各 RQ に対し考察を行う。

5.1 RQ1

以下に各トピックごとの詳細および各トピックの中で引用数が多い上位 2 本の論文の概要を示す。

5.1.1 FP の利点や欠点の評価

本トピックに含まれる研究の内訳について表 16 に示す。

本トピックには、主に FP そのものの評価についての研究が含まれている。FP の計測精度や信頼性、利用した上で得られる利益といった情報は FP を利用か判断する上で非常に重要な基準であり、FP が提案されて以降研究が行われてきた。計測精度の評価としては「他の開発規模尺度との比較」[34, 37, 47, 81] や「FP 法同士の比較」[43, 85, 96, 122], 「早期見積もりにおいて FP を利用した際の計測精度の評価」[44] などが行われ、FP の利益 [39, 57, 70, 105] や信頼性 [42, 94], 問題点 [48, 51] についての評価も行われた。

その他の評価研究としては、重みづけ [50] や調整係数 [66], 計測要素 [36, 45, 54, 112] といった計測項目が評価され、開発規模を計測する上で正しく機能しているかを確認されたほか、LOC との関係性の評価 [32, 59, 99, 109] や FP の適切な使用法のと環境の調査 [68] など FP の利用にまつまる調査も行われた。

5.1.2 計測補助

本トピックに含まれる研究の内訳について表 17 に示す。

本トピックには、主に FP 計測の補助についての研究が含まれている。FP の利用及び普及における大きな障壁として計測の手間が挙げられる。行数を測るだけでよい LOC 等と比べ、計測要素の多い FP を算出するためには非常に時間がかかる。これに加えて、基本的に自由記述である仕様書から FP を計測するためには主観的判断を含める必要があり、そのため FP の計測結果には計測者によって誤差が生じると言われている。

上記の問題に対し、UML[72, 78, 82, 83, 84, 86, 88, 89, 104, 113, 115, 124] やソースコード [69, 102, 123] などからの計測およびそれらの自動化が研究されている。自動計測によって計測の手間は大きく削減でき、たとえ自動計測でないとしても、記述方法が定められている UML モデルなどからの計測は仕様書からの計測に加え主観的判断を行う機会が減少し、結果として計測コストや誤差の削減に繋がる。また、ソースコードからの自動計測は開発初期

表 16: 「評価」トピックの研究内訳

小トピック	対象論文	本数
他の開発規模尺度との比較	[34, 37, 47, 81]	4
LOC との関係の評価	[32, 59, 99, 109]	4
FP 法同士の比較	[43, 85, 96, 122]	4
FP の利益の評価	[39, 57, 70, 105]	4
計測要素の評価	[36, 45, 54, 112]	4
FP の問題点の評価	[48, 51]	2
信頼性評価	[42, 94]	2
重みづけの評価	[50]	1
調整係数の評価	[66]	1
早期見積もりにおける評価	[44]	1
適切な使用方法と環境の調査	[68]	1

の計測には利用できないものの、仕様書などが残っておらずソースコードしか現存していない既存プロジェクトの開発規模を計測する上で非常に有用である。UML やソースコードのほかに、自然言語 [106] やフォーマット [121] への入力からの自動計測についても研究されている。

5.1.3 計測された FP の活用

本トピックに含まれる研究の内訳を表 18 に示す。

本トピックには、主に計測された FP の活用についての研究が含まれている。FP の活用用途としては主に「工数の導出」と「定量的評価」の 2 つがある。FP はあくまで開発規模の尺度であり、予算等を算出するには FP を時間などの工数に変換する必要がある。しかしながら、一方で FP と工数の関係は非常に複雑であり、一般的な変換モデルが存在していない。そのため、以前からその変換手法について研究が進められており [49, 55, 56, 61, 91, 98, 103, 108, 110], FP から直接予算を見積もるための研究も行われている [117]。

また、FP はソフトウェアの機能を元にその規模を計測するため、Lines of Code(LOC) と違って計測結果が開発言語に依存しない。そこに着目し、人月あたりの LOC に代わって人月あたりの FP を算出する事で生産性を評価する研究が行われている [33, 65]。LOC に対して優位な点としては、「異なる開発言語であっても比較が可能」「実装の際個人によってばらつきが出る LOC と違い FP では大きなばらつきが生じない」等の点がある。FP による定

表 17: 「計測補助」トピックの研究内訳

小トピック	対象論文	本数
UML からの計測	[72, 78, 82, 83, 84, 86, 88, 89, 104, 113, 115, 124]	12
ソースコードからの計測	[69, 102, 123]	3
DFD モデルからの計測	[58, 73]	2
OO-H モデルからの計測	[101, 125]	2
自然言語からの計測	[106]	1
フォーマットへの入力による計測	[121]	1
その他モデルからの計測	[60, 74, 77, 97]	4

表 18: 「活用」トピックの研究内訳

小トピック	対象論文	本数
工数	[49, 55, 56, 61, 91, 98, 103, 108, 110]	9
生産性	[33, 65]	2
開発言語の水準評価	[52]	1
スケジュールの定量評価	[53]	1
価格設定	[117]	1

量的評価においては、1 行あたりの FP を算出する事で開発言語の水準を定量的の評価する試み [52] や、スケジュールを定量的に評価する試み [53] が行われている。

5.1.4 計測ルールの変更

本トピックに含まれる研究の内訳について表 19 に示す。

本トピックには、主に FP の計測ルールの是正についての研究が含まれている。FP の計測ルールは専門家による話し合いによって決定されたものであり、その重みづけなどに理論的根拠が存在しない [50]。そのため、その正確さが問題視されており、是正のための研究が行われている。研究されているのは主に重みづけと計測要素、そして調整係数についてである。重みづけにおいてはその値や範囲設定が問題視されており、値や範囲の再設定について研究が進められている [41, 46, 63, 76, 90]。計測要素においては「全ての計測要素が必要なのか否か」、つまりより少ない計測要素でも開発規模の計測が可能ではないかと指摘されて

表 19: 「ルール変更」トピックの研究内訳

小トピック	対象論文	本数
重みづけの変更	[41, 46, 63, 76, 90]	5
計測要素の変更	[64, 111]	2
調整係数の変更	[40]	1
ルールの追加	[79, 100, 118]	3
新ルールの提案	[35, 67, 80]	2

おり、計測要素と工数の相関や計測要素の削減について研究が行われている [64, 111]。調整係数については「取り扱っている技術的要因が現代に合っていない」という指摘があり、新たな技術的要因の設定について研究が行われている [39, 46]。また、上記の研究のように計測ルールの一部のみを変更するのではなく、ルールの追加 [79, 100, 118] やルール全体の変更、つまり新たな計測ルールの提案 [35, 67, 80] を行っている研究も存在する。

5.1.5 計測が難しいとされている対象への適用

本トピックに含まれる研究の内訳について表 20 に示す。

本トピックには、主に特定のソフトウェアや開発手法への FP の適用についての研究が含まれている。FP はユーザ視点で機能を計測するためにその内部処理を考慮する必要がない反面、リアルタイムシステムや組込みシステム等、内部処理が複雑なシステムに対しては正確に開発規模が計測できない事を指摘されている。また、FP の計測のためには基本的にシステム全体の要件が必要であり、要件を最初に全て決定するウォーターフォール型開発では問題なく扱えるものの、要件が初期に全て揃わないアジャイル開発などの非ウォーターフォール型開発では扱う事が難しい。加えて、システム保守における開発規模の計測も難しいとされている。

上記のソフトウェアや開発手法に対し FP を適用するため、適用方法の提案などの研究が行われている。非ウォーターフォール型開発に対する FP の適用としては、インクリメンタル型開発 [114] や反復型開発 [93] に対する FP の適用についての研究が行われており、ソフトウェア保守についても FP 適用についての研究が行われている [62, 119, 120]。特定のソフトウェアに対する FP の適用については、データマートシステム [71] や組込みシステム [95]、リアルタイムシステム [38, 95] やバーチャルリアリティシステム [75] への適用についての研究が進められている。

表 20: 「適用」トピックの研究内訳

小トピック	対象論文	本数
ソフトウェア保守への適用	[62, 119, 120]	3
非ウォータフォールモデルへの適用	[93, 114]	2
リアルタイムシステムへの適用	[38, 95]	2
データマートシステムへの適用	[71]	1
科学システムへの適用	[38]	1
組込みシステムへの適用	[95]	1
バーチャルリアリティシステムへの適用	[75]	1

表 21: 「その他」トピックの研究内訳

小トピック	対象論文	本数
IFPUG 法から COSMIC 法への変換	[87, 92, 107, 126]	4
FP の問題点とその改善についての SLR	[116]	1

5.1.6 その他

本トピックに含まれる研究の内訳について表 21 に示す。

本トピックには、前述の 5 つのトピックに含まれなかった研究が含まれている。内訳としては「IFPUG 法から COSMIC 法への変換についての研究」と「系統的文献レビュー」である。

主流な FP 法として IFPUG 法と COSMIC 法が存在する。双方がソフトウェアの開発規模の尺度であり、それぞれ IFPUG-FP と COSMIC-FP を計測する。しかしながら、双方の FP は計測手法が異なるため一致せず、その点において問題が生じている。双方の FP の比較や、IFPUG 法で得た知見や蓄積データを COSMIC 法に適用するためには IFPUG 法で得た FP を COSMIC 法の FP に変換する必要がある、変換手法についての研究が行われている。[87, 92, 107, 126]

また、本トピックには「FP の問題点と改善点についての系統的レビュー」[116] も含まれている。この研究は FP の計測上の問題とその改善策についての論文を 2002 年から 2014 年の範囲で収集し、結果集まった 18 本の論文における研究成果を整理・分析したものである。

5.2 RQ2

各トピックの研究動向の考察を行う。

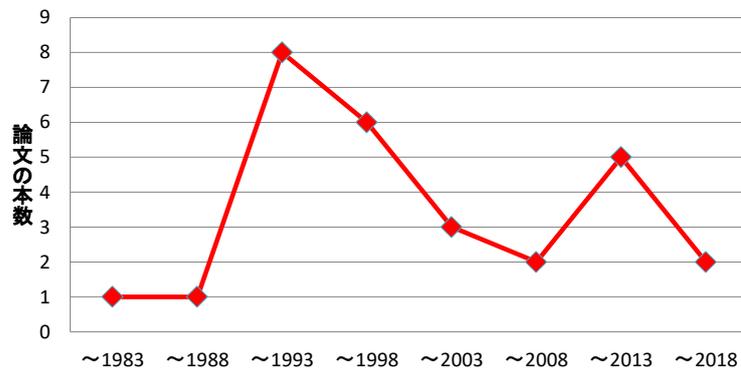


図 5: 出版された年代ごとの「評価」トピックの論文数

5.2.1 「評価」トピックの研究動向

出版年代ごとの論文数を図 5 に示す。

本トピックにおいては、1989 年から 1993 年において特に研究が盛んとなっている。1989 年から 1993 年においては、後述の「活用」トピックや「ルール変更」トピックと同じく、この時期に FP が十分に普及したためその計測精度や信頼性、他手法との比較に注目が集まったと考えられる。研究内容としても FP 法の利益 [39] や信頼性 [42]、計測要素 [35, 44] の評価のほか、他開発規模尺度との比較 [36, 45] や早期見積もりにおいて FP を利用した際の計測精度の評価 [44] など多岐に渡っており、FP に対する関心が非常に高まっていた事が見受けられる。

1994 年以降の「評価」トピックの研究数は減少しているものの、各年代において 1 本は研究が行われており、FP に対する関心が薄れていない事が伺える。特に 1994 年以降で最も本数が多い年代である 2009 年から 2013 年においては 5 本中 3 本が FP の 1 つである COSMIC 法に関する論文 [96, 99, 109] であり、後述する「その他」トピックと合わせて COSMIC 法に対する関心が強まっていると言える。一方で、COSMIC 法の計測精度や信頼性についての研究は未だ少なく、利用者が COSMIC を選択する上での判断基準となる情報が乏しい現状にある。そのため、「評価」トピックにおいては COSMIC 法の評価、計測精度や信頼性および IFPUG 法などの他 FP 法との比較についての研究が行われるべきだと言える。

5.2.2 「計測補助」トピックの研究動向

出版年代ごとの論文数を図 6 に示す。

本トピックは、1990 年代に盛んであった「活用」トピックや「ルール変更」トピックと異なり、2000 年以降に研究が発展している。これは研究の大部分を占める UML が 1997 年に考案され、そこから普及が始まったためであると考えられる。研究が最も多い 2004 年から

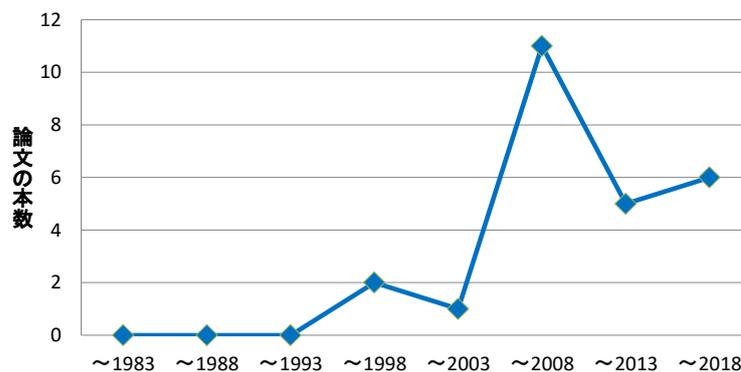


図 6: 出版された年代ごとの「計測補助」トピックの論文数

2008年における研究は11本中8本がUMLからのFP計測 [72, 78, 82, 83, 84, 86, 88, 89] についてであり、研究者がUMLからのFP計測に対し多大な期待を寄せている事が見て取れる。また、本トピックでは2004年以降5年ごとに5本以上の研究が行われており、他のトピックと比べても安定したものとなっている。これはFP計測の手間が解決すべき課題として今日に至るまで捉えられている事を示唆しており、FP利用の簡便化および更なる普及への助力のために今後も研究を継続する必要があるトピックであると言える。

また、今回対象とした論文の中では1本のみであるが、自然言語で記述された仕様書からのFP計測 [106] が研究されており、このような研究は自然言語に対する解析技術が向上するにつれてより発展するものと考えられるため、今後の技術進歩によっては更なる発展が期待される。

5.2.3 「活用」トピックの研究動向

出版年代ごとの論文数を図7に示す。

このトピックにおいては、1994年から1998年に研究が最も盛んとなっている。これは1979年に考案されたFP法が十分に普及し、その活用に注目が寄せられたためと考えられる。研究内容としても、工数 [49, 55, 56, 61] だけでなく開発言語の水準としてのFPの活用 [52] やスケジュールの定量的評価へのFPの活用 [53] など、FPの活用に対し意欲的に研究されているのが見て取れる。

2000年以降においても研究は行われており、特に2009年から2014年における論文数は1994年から1998年の次に多い。2000年以降の研究において7本中5本が工数についての研究である事から、FPを用いた工数見積もりへの関心は今もなお強い物であると考えられる。また、2000年以降の「活用」トピックの研究は「開発早期の計画や要求分析にかかる工数からの開発全体の工数の見積もり」 [108] とこれまでになく着眼点での工数導出や、

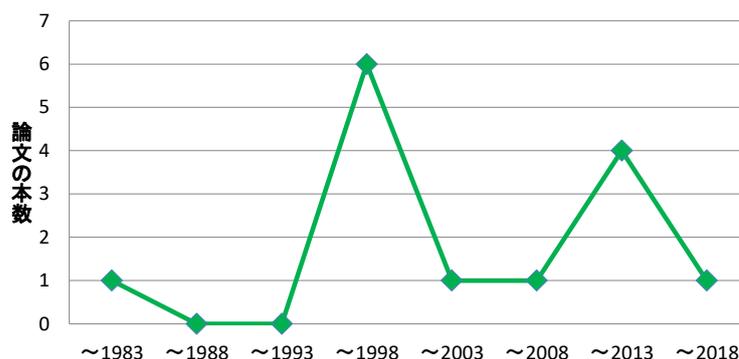


図 7: 出版された年代ごとの「活用」トピックの論文数

「ニューラルネットワークを用いた FP と開発環境からの工数見積もり」[91, 103, 110] と新技術を用いた工数導出が行われている。特に後者は、今後情報科学分野において新技術が生じる度に FP からの工数導出についての研究が進展する可能性を示唆しており、研究者は新技術が FP からの工数導出に適用できるか否かを常に実践する必要があると言える。

5.2.4 「ルール変更」トピックの研究動向

出版年代ごとの論文数を図 8 に示す。

本トピックにおいては、1989 年から 1993 年と 2004 年から 2008 年において研究が盛んとなっている。1989 年から 1993 年においては前述の「活用」トピックと同じく、FP の普及が十分に進んだため FP の重みづけなどに疑問が抱かれるようになったと考えられる。2004 年から 2008 年においては「ニューラルネットワークを用いた重みづけの値の再設定」[76] や「統計解析とニューラルネットワークとファジィ理論を用いた重みづけの較正」[90] と「活用」トピックと同様に新たな技術であるニューラルネットワークによって研究が進められている。そのため、「ルール変更」分野も「活用」トピックと同様に今後情報科学分野における技術発展によって研究が進展する可能性があり、研究者が新技術の適用を常に実践する必要がある分野であると言える。

5.2.5 「適用」トピックの研究動向

出版年代ごとの論文数を図 9 に示す。

本トピックにおいては、2004 年以降研究が盛んであり、直近 15 年間における各トピックの本数を比較すると表 22 に示した通り「適用」トピックが 3 番目に多い。これはここ 15 年においてソフトウェアや開発手法が多様化し、これまでの適用方法では計測が難しくなったためと考えられる。今後もソフトウェアや開発手法の多様化は進むと考えられるため、必然

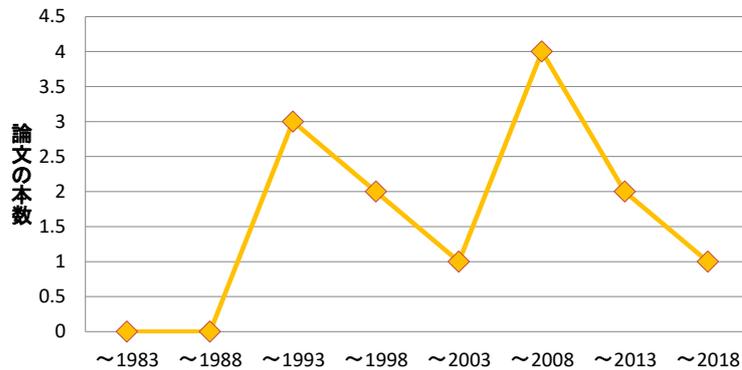


図 8: 出版された年代ごとの「ルール変更」トピックの論文数

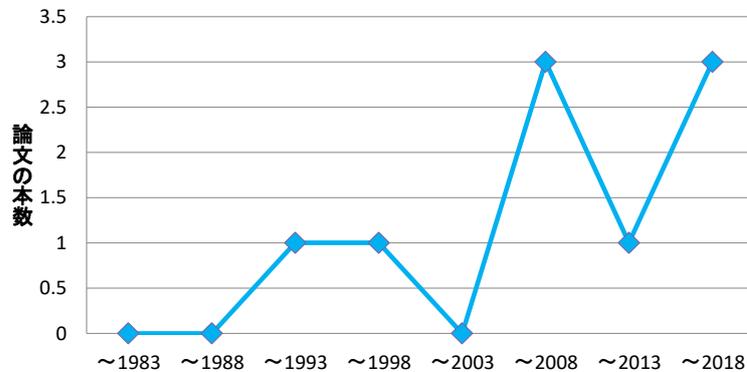


図 9: 出版された年代ごとの「適用」トピックの論文数

的にこのトピックにおける研究もまた、継続して行われるべきであると言える。

5.2.6 「その他トピック」の研究動向

出版年代ごとの論文数を図 10 に示す。

IFPUG 法から COSMIC 法への変換については 2004 年以降から研究が行われており、この 15 年で 4 本研究されている。この本数は今回対象とした論文 95 本において小トピックで順位付けした場合、表 23 に示した通り 3 位となる程大きな割合を占めるものであり、非常に盛んに研究されている分野であると言える。そのため、研究者の COSMIC 法への関心は非常に強い物であると考えられる。加えて、COSMIC 法の利用を判断するための基準の 1 つとして、IFPUG 法からの COSMIC 法への変換が可能か否かという情報は過去の蓄積データや知見の利用に関わる重要なものであるため、今後も研究する必要があると言える。

また、FP に関する SLR は FP が考案された 1979 年から 2018 年現在において「FP の問題点と改善点についての系統的レビュー」[116]のみであった。研究成果の整理・分析および今

表 22: 直近 15 年における各トピックの論文本数

トピック	本数
計測補助	22
評価	9
適用	7
ルール変更	7
活用	6
その他	4

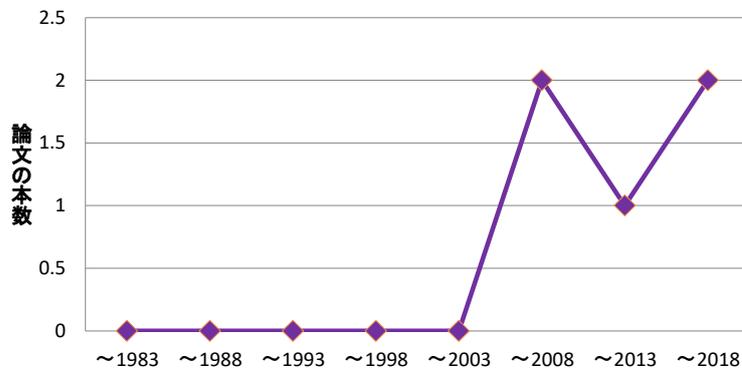


図 10: 出版された年代ごとの「その他」トピックの論文数

後の研究課題の特定のために、今後 FP の様々な分野に対する SLR が実施されるべきである。

5.3 RQ3

5.3.1 研究が適用可能なコンテキスト情報

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持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についての情報が有用であるかが不明であるため、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の特定も今後の研究課題として考えられる。

表 23: 直近 15 年における論文本数の多い小トピック (上位 3 つ)

トピック	小トピック	本数
計測補助	UML からの計測	12
活用	工数	5
その他	IFPUG 法から COSMIC 法への変換	4

5.3.2 検証に用いるプロジェクト

Observational Methods および Historical Methods 内の Study of Legacy において、検証に用いられるプロジェクト数が少ないために妥当性が十分示されておらず、そのため開発現場への適用可能性が保証されているとは言えない研究が存在した。また、Observational Methods においては検証に用いるプロジェクトが単一の企業の物であるために他企業における研究成果の妥当性が示されておらず、そのため開発現場への適用可能性が保証されていないと言える研究が存在した。検証に用いられるプロジェクト数が少ない事や単一の企業からしかプロジェクトの提供を受けていない原因として、FP 研究の検証のために必要なデータが企業から得られ辛いという現状がある。FP の計測のためには要求仕様書等のソフトウェアの仕様について記載された文書が必要となる。しかしながら、大半の企業においてこの類の文書は発注者との守秘義務契約により研究への提供が難しい事が多く、そのため上述のような問題が生じている。表 9 に記載されている ISBSG データセットは様々な国や企業のプロジェクトのデータが収集されているため提案手法の検証等に有用ではあるが、このデータセットには要求仕様書等が含まれておらず、また、検証したい研究に適したデータが存在していない可能性もあるため、FP についての全ての研究を ISBSG データセットで検証する事は不可能である。この問題を解決するために、要求仕様書等が蓄積された大規模な検証用のデータセットの作成が考えられる。ISBSG データセットのように様々な国や企業のプロジェクトにおける仕様書や実際にかかった工数などが蓄積されたデータセットが存在すれば、FP 分野における研究成果の検証はより多くのプロジェクトを用いて行う事ができるようになり、その結果開発現場での適用を考慮するに値する程の妥当性を示す事が可能となる。データセット作成の源としては、官公庁などが実施している公的なプロジェクトの取捨が考えられる。これは守秘義務などにより研究者への公開が難しい一般企業と比べ、官公庁などのプロジェクトは公共事業であるため公開が比較的容易であると考えられるためである。

5.4 RQ4

5.4.1 汎用的な自動計測ツールがない

人の手によるFPの計測には手間がかかるため、自動計測ツールの実現が強く望まれている。自動計測ツールについてはUMLからの自動計測など15本の研究が行われていた。しかしながら、IFPUGから公認を受けた自動計測手法[23]は現在存在しておらず、開発現場での適用は未だ難しいと言える。そのため、今後も自動計測ツールについての研究は継続して行われるべきである。また、提案されている自動計測手法は主にUMLのようなモデルから計測を行うため、自動計測ツールを導入する場合は計測元となるUML等のモデルも導入する必要があり、UMLを用いていない企業にとっては導入コストがかかる。そのため、今後は自然語で記述された要求仕様書からの自動計測などのような、モデルの導入コストがかからない手法についても研究が行われるべきである。

5.4.2 保守案件への適用が難しい

保守案件の工数を見積もるためにFPを適用する研究は複数行われているものの十分な精度を得られていない。その原因として、保守案件の工数の特殊性が考えられる。保守、つまりシステムの変更を行う場合、その工数は変更の規模だけでなくソフトウェアの品質に大きく影響を受ける。ソースコードを例にすると、ソースコードが複雑に記載されておりなおかつコメントなども残されていない場合、変更の前にソースコードそのものを読み解く手間などが生じるため、変更にかかる工数は増大する。反対に、ソースコードが簡潔に記載されており、記載内容の理解が容易である場合、ソースコードを解釈する手間が省略される分変更にかかる工数は少なくなる。ソフトウェアの品質によって工数が増減する以上、機能規模であるFPのみによる工数見積もりは難しいと言える。そのため、今後の研究課題として「変更するソフトウェアの品質を考慮に入れた工数見積もり」が挙げられる。ISO/IEC 25010:2011[24]の保守性などを参考にソフトウェアの品質についての指標を定義し、それとFPを元に工数見積もりを行う事により、保守案件における工数見積もりを行う事が考えられる。

5.4.3 生産性評価モデルが確立されていない

FPを用いた生産性評価モデルとして、時間/FPというものがある。このモデルは1FP当たりにかかった時間によって生産性を評価するものであり、容易に算出が行える事もあってFPが提案されて以降用いられてきた評価指標である。しかしながら、開発規模であるFPのみでは生産性を正確に評価する事は難しく[65]、それ以外の要因を加えた非公式なモデルが

様々な企業で考案されている。そういった背景の元、開発現場のニーズとしてより正確な生産性評価モデルが求められているが、研究本数は2本と少なく、未発達な研究分野であると言える。そのため、より正確な生産性評価のためにはどのような説明変数が有効かについて、今後も継続して研究が行われるべきである。また、生産性評価についての研究が行われていない原因として、必要な情報の多さが挙げられる。生産性評価のための新たな計測要因の模索のためにはプロジェクトにおける様々な情報が必要となるが、前述の通り企業のプロジェクトは手に入り辛いために研究が実行し難いと考えられる。そのため、生産性評価モデルの研究をより促進するためには、プロジェクトについての様々な情報が蓄積された大規模なデータセットの作成が必要であると言える。

5.4.4 発注側企業にFPが普及していない

FPを見積もりの根拠として扱う場合、開発側だけでなく発注側もFPを理解している必要がある。しかしながら、現状において発注側にはFPが十分普及しておらず[30]、原因としてFPを利用した際のユーザ側のメリットが不明であるために普及が進んでいない事が挙げられた。FPを用いた際の開発側のメリットとしては「ソースコードの行数などの他の開発規模の尺度と比較して工数の計測精度が高い」「客観的な根拠が高い」等が周知されており、既存研究においても開発側にとってのメリットを調査する研究は存在している[39, 57, 70, 105]。一方で、発注側のメリットは周知されておらず、研究も95本中1本[117]しか行われていなかった。そのため、FPの普及のために発注側のメリットを評価する研究が必要であると言える。

5.4.5 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、非機能要件を工数見積もりなどに反映させる事が不可能という問題がある。FP提案当初と比べ非機能要件の比率が増加の一途を辿っている事もあり、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もりについての研究[110]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、工数見積も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もりはそうでないものと比べて高精度」「FPとデータベースの種類を入力とした見積もりが最も高精度」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もりは非機能要件を組み合わせる事で見積もり精度が向上される事が分かった。そのため、今後、FPと組み合わせる事で工数見積も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。

5.4.6 FPのよい教育手法がない

FPの教育手法についての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ため、今後研究が必要な課題である。FP計測において重要な点は「計測ルールを理解」と「情報の補完」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と考えられる。また、要求仕様などからFPを計測する場合、情報の補完が必要な状況が存在する。例を挙げると、要求仕様において1つの機能として記載されているものが実際のソフトウェアにおいては複数の機能として実装される事があり、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。その際計測者は要求仕様に記載されている情報に加え業務知識等を元に自ら情報を補完して計測を行う必要があるが、FP計測の経験が浅い者が計測者である場合補完が行えない事がある。それを解決するためには、FP計測の実績を蓄積し共有する事で情報の補完方法を学習する必要がある。よって、今後行われるべき研究としては「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」および「情報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。

6 あとがき

本章では、本研究のまとめと今後の課題について述べる。

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 6 つのトピックに分類し、それぞれのトピックにおいて研究内容の整理と研究動向の分析を行った。その結果、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 の工数見積も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、研究成果が開発現場で適用可能かどうかを調査するために、研究成果が適用可能となるプロジェクトのコンテキストについて詳細に記載が行われているかを調べた結果、大半の研究においてコンテキストについての記載が行われておらず開発現場への適用を考える上での判断材料が不足している事が分かった。また、開発現場のニーズと研究にとのギャップの有無を調査するために開発現場とのニーズと既存研究を照らし合わせた結果、現場のニーズがあるにもかかわらず教育手法についての研究が行われていない事や自動計測手法などの研究が現場のニーズを現状満たしていない事が分かった。

今後の課題として、「教育手法の提案」や「発注側企業への FP の普及」、「検証用データセットの構築」などといった FP 研究における今後の課題への取り組みが挙げられる。

謝辞

本研究の全過程を通し，理解ある親身なご指導を賜り，的確なご助言を頂きました楠本 真二教授に心より感謝申し上げます。

本研究を行うにあたり，的確なご指導およびご助言を頂きました肥後 芳樹 准教授に心より感謝申し上げます。

本研究を行うにあたり，丁寧かつ適切なお指導を頂きました 松本 真佑 助教に深く感謝申し上げます。

本研究を進めるにあたって，貴重なご助言を頂きました株式会社 SOFTEST 倉重誠氏，一般財団法人経済調査会 大岩佐和子氏，押野智樹氏，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。

研究室生活の中で相談に乗って頂き，また励まして頂きました大阪大学大学院情報科学研究科コンピュータサイエンス専攻博士前期課程2年の有馬諒氏，同佐々木美和氏，同谷門照斗氏，同内藤圭吾氏，同松尾裕幸氏に深く感謝申し上げます。

研究室生活を大変豊かにして頂きました大阪大学大学院情報科学研究科コンピュータサイエンス専攻博士前期課程1年の田中紘都氏，同土居真之氏，同中島望氏，同林純一氏，同松本淳之介氏に深く感謝申し上げます。

研究室の環境維持に多くのご助力を頂きました，大阪大学基礎工学部情報科学科4年のキム テヨン氏，同富田裕也氏，同中川将氏，同華山魁生氏に深く感謝申し上げます。

最後に，本研究に至るまでに，講義，演習，実験等でお世話になりました大阪大学基礎工学部情報科学科の諸先生方に，この場を借りて心から御礼申し上げます。

参考文献

- [1] B. Boehm, C. Abts, S. Chulani. Software development cost estimation approaches - A survey. *Annals of software engineering*, Vol. 10, No. 1-4, pp. 177-205, Springer, 2000.
- [2] R. L. Glass. *Facts and fallacies of software engineering*, Addison-Wesley Professional, 2002.
- [3] A. J. Albrecht. Function point analysis. *Encyclopedia of Software Engineering*, Vol. 1, pp. 518-524, Addison-Wesley Professional, 1994.
- [4] 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.
- [5] IFPUG: Function Point Counting Practices Manual, Release 4. 3. International Function Point Users Group, 2010.
- [6] C. R. Symons. *Software sizing and estimating: Mk II FPA (function point analysis)*, John Wiley & Sons, Inc. , 1991.
- [7] 山口正明 (著) 調重俊 (監修). ファンクションポイント cosmic- ffp 法実践ガイドー組込み系・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.
- [8] 柏本隆志, 楠本真二, 井上克郎, 鈴木文音, 湯浦克彦, 津田道夫. イベントトレース図に基づく要求仕様書からのファンクションポイント計測手法. Vol. 41, No. 6, pp. 1895-1904, 2000.
- [9] Barbara Kitchenham. Procedures for performing systematic reviews. Keele, UK, Keele University, Vol. 33, No. 2004, pp. 1-26, 2004.
- [10] Pearl Brereton, Barbara A Kitchenham, David Budgen, Mark Turner, and Mohamed Khalil. Lessons from applying the systematic literature review process within the software engineering domain. *Journal of systems and software*, Vol. 80, No. 4, pp. 571-583, 2007.
- [11] Barbara Kitchenham, Pearl Brereton, Zhi Li, David Budgen, and Andrew Burn. Repeatability of systematic literature reviews. In *Evaluation & Assessment in Software Engineering (EASE 2011)*, 15th Annual Conference on, pp. 46-55, 2011.

- [12] Barbara Kitchenham and Pearl Brereton. A systematic review of systematic review process research in software engineering. *Information and Software Technology*, Vol. 55, No. 12, pp. 2049-2075, 2013.
- [13] Jianfeng Wen, Shixian Li, Zhiyong Lin, Yong Hu, and Changqin Huang. Systematic literature review of machine learning based software development effort estimation models. *Information and Software Technology*, Vol. 54, No. 1, pp. 41-59, 2012.
- [14] Ali Idri, Fatima azzahra Amazal, and Alain Abran. Analogy-based software development effort estimation: A systematic mapping and review. *Information and Software Technology*, Vol. 58, pp. 206-230, 2015.
- [15] Computing Research & Education; CORE Rankings Portal. <http://www.core.edu.au/index.php/conference-portal>
- [16] Hirsch, Jorge E. "An index to quantify an individual's scientific research output." *Proceedings of the National academy of Sciences* 102.46 (2005): 16569-16572.
- [17] Magne Jorgensen and Martin Shepperd. A systematic review of software development cost estimation studies. *Software Engineering, IEEE Transactions on*, Vol. 33, No. 1, pp. 33-53, 2007.
- [18] IEEE xplore. <http://ieeexplore.ieee.org/Xplore/home.jsp>.
- [19] ACM Digital library. <http://dl.acm.org/>.
- [20] Science Direct. <http://www.sciencedirect.com/>.
- [21] Scopus <https://www.scopus.com/>
- [22] Google Scholar. <https://scholar.google.co.jp/>.
- [23] IFPUG SOFTWARE CERTIFICATION <http://www.ifpug.org/certification/software-certification/>
- [24] システム及びソフトウェア製品の品質要求及び評価 (SQuaRE)-システム及びソフトウェア品質モデル <http://kikakurui.com/x25/X25010-2013-01.html>
- [25] Zelkowitz, Marvin V., and Dolores R. Wallace. "Experimental models for validating technology." *Computer* 31.5 (1998): 23-31.

- [26] Sturges, Herbert A. "The choice of a class interval." *Journal of the american statistical association* 21.153 (1926): 65-66.
- [27] Ivanov, Vladimir, et al. "What do software engineers care about? gaps between research and practice." *Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering*. ACM, 2017.
- [28] ABOUT THE XML METADATA INTERCHANGE SPECIFICATION VERSION 2.0 <https://www.omg.org/spec/XMI/2.0/About-XMI/>
- [29] Ambler, Scott W. *The object primer: Agile model-driven development with UML 2.0*. Cambridge University Press, 2004.
- [30] ユーザー企業ソフトウェアメトリックス調査 <http://www.juas.or.jp/cms/media/2017/02/14swm.pdf>
- [31] ソフトウェア開発データ白書 2016-2017 <https://www.ipa.go.jp/files/000057877.pdf>
- [32] Albrecht, Allan J., and John E. Gaffney. "Software function, source lines of code, and development effort prediction: a software science validation." *IEEE transactions on software engineering* 6 (1983): 639-648.
- [33] Behrens, Charles A. "Measuring the productivity of computer systems development activities with function points." *IEEE Transactions on Software Engineering* 6 (1983): 648-652.
- [34] Kemerer, Chris F. "An empirical validation of software cost estimation models." *Communications of the ACM* 30.5 (1987): 416-429.
- [35] Symons, Charles R. "Function point analysis: difficulties and improvements." *IEEE transactions on software engineering* 14.1 (1988): 2-11.
- [36] Verner, June M., et al. "Technology dependence in function point analysis: a case study and critical review." *Software Engineering*, 1989. 11th International Conference on. IEEE, 1989.
- [37] Low, Graham C., and D. Ross Jeffery. "Function points in the estimation and evaluation of the software process." *IEEE Transactions on Software Engineering* 1 (1990): 64-71.
- [38] Reifer, Donald J. "Asset-R: A function point sizing tool for scientific and real-time systems." *Journal of Systems and Software* 11.3 (1990): 159-171.
- [39] Heemstra, F. J., and R. J. Kusters. "Function point analysis: Evaluation of a software cost estimation model." *European Journal of Information Systems* 1.4 (1991): 229-237.

- [40] Guerra, L. G., and H. S. Lobos. "An Empirical Validation of a Software Cost Model." IFAC Proceedings Volumes 23.5 (1990): 57-62.
- [41] Bock, Douglas B., and Robert Klepper. "FP-S: A simplified function point counting method." *Journal of Systems and Software* 18.3 (1992): 245-254.
- [42] Kemerer, Chris F., and Benjamin S. Porter. "Improving the reliability of function point measurement: an empirical study." *IEEE Transactions on Software Engineering* 18.11 (1992): 1011-1024.
- [43] Rask, Raimo, Petteri Laamanen, and Kalle Lyytinen. "A comparison of Albrecht's function point and Symons' Mark II metrics." ICIS. 1992.
- [44] Betteridge, R. "Successful experience of using function points to estimate project costs early in the life-cycle." *Information and Software Technology* 34.10 (1992): 655-658.
- [45] Kitchenham, Barbara, and Kari Kansala. "Inter-item correlations among function points." *Software Metrics Symposium, 1993. Proceedings., First International.* IEEE, 1993.
- [46] Jeffery, D. Ross, Graham C. Low, and M. Barnes. "A comparison of function point counting techniques." *IEEE Transactions on Software Engineering* 19.5 (1993): 529-532.
- [47] Rask, Raimo, Petteri Laamanen, and K. Lyytinen. "Simulation and comparison of Albrecht's function point and DeMarco's function bang metrics in a CASE environment." *IEEE Transactions on Software Engineering* 19.7 (1993): 661-671.
- [48] Hughes, R. "An exploration of the problems of using function points mark II." *Information Systems Journal* 4.3 (1994): 169-183.
- [49] Matson, Jack E., Bruce E. Barrett, and Joseph M. Mellichamp. "Software development cost estimation using function points." *IEEE Transactions on Software Engineering* 20.4 (1994): 275-287.
- [50] Abran, Alain, and Pierre N. Robillard. "Function points: a study of their measurement processes and scale transformations." *Journal of Systems and Software* 25.2 (1994): 171-184.
- [51] Kitchenham, Barbara, Shari Lawrence Pfleeger, and Norman Fenton. "Towards a framework for software measurement validation." *IEEE Transactions on software Engineering* 21.12 (1995): 929-944.

- [52] Jones, Capers. "Backfiring: Converting lines of code to function points." *Computer* 28.11 (1995): 87-88.
- [53] Jones, Capers. "Determining software schedules." *Computer* 28.2 (1995): 73-75.
- [54] Abran, Alain, and Pierre N. Robillard. "Function points analysis: An empirical study of its measurement processes." *IEEE Transactions on Software Engineering* 22.12 (1996): 895-910.
- [55] Subramanian, Girish H., and George E. Zarnich. "An examination of some software development effort and productivity determinants in ICASE tool projects." *Journal of Management Information Systems* 12.4 (1996): 143-160.
- [56] Finnie, Gavin R., Gerhard E. Wittig, and Jean-Marc Desharnais. "A comparison of software effort estimation techniques: using function points with neural networks, case-based reasoning and regression models." *Journal of systems and software* 39.3 (1997): 281-289.
- [57] Furey, Sean. "Why we should use function points [software metrics]." *IEEE Software* 14.2 (1997): 28.
- [58] Shoal, Peretz, and Opher 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* 39 (1997): 855-865.
- [59] Dolado, Jose Javier. "A study of the relationships among Albrecht and Mark II function points, lines of code 4GL and effort." *Journal of Systems and Software* 37.2 (1997): 161-173.
- [60] Antoniol, Giuliano, et al. "Adapting function points to object oriented information systems." *International Conference on Advanced Information Systems Engineering*. Springer, Berlin, Heidelberg, 1998.
- [61] Fujiwara, Fumiko, Takushi Goto, and Sadao Araki. "Examples of applying software estimate tool." *Software Engineering, 1998. Proceedings of the 1998 International Conference on*. IEEE, 1998.
- [62] Fleck, Robert A., and Jr. "Managing programmer resources in a maintenance environment with function points." *Industrial Management & Data Systems* 98.2 (1998): 63-70.

- [63] Yau, Chuk, and Ho-Leung Tsoi. "Modelling the probabilistic behaviour of function point analysis." *Information and Software Technology* 40.2 (1998): 59-68.
- [64] Horgan, Gerard, Souheil Khaddaj, and Peter Forte. "Construction of an FPA-type metric for early lifecycle estimation." *Information and Software Technology* 40.8 (1998): 409-415.
- [65] Bok, Hai Suan, and Krishnamurthy S. Raman. "Software engineering productivity measurement using function points: a case study." *Journal of Information Technology* 15.1 (2000): 79-90.
- [66] Lokan, Christopher J. "An empirical analysis of function point adjustment factors." *Information and Software Technology* 42.9 (2000): 649-659.
- [67] Orr, George, and Thomas E. Reeves. "Function point counting: one program 's experience." *Journal of Systems and Software* 53.3 (2000): 239-244.
- [68] Van Genuchten, Michiel, and Hans Koolen. "On the use of software cost models." *Information & Management* 21.1 (1991): 37-44.
- [69] Kusumoto, Shinji, et al. "Function point measurement from Java programs." *Proceedings of the 24th International Conference on Software Engineering*. ACM, 2002.
- [70] Kitchenham, Barbara, et al. "An empirical study of maintenance and development estimation accuracy." *Journal of systems and software* 64.1 (2002): 57-77.
- [71] Calazans, Angelica Toffano, Kathia Marcal de Oliveira, and Rildo Ribeiro dos Santos. "Adapting function point analysis to estimate data mart size." *Software Metrics, 2004. Proceedings. 10th International Symposium on*. IEEE, 2004.
- [72] Cantone, Giovanni, Davide Pace, and Giuseppe Calavaro. "Applying function point to unified modeling language: Conversion model and pilot study." *Software Metrics, 2004. Proceedings. 10th International Symposium on*. IEEE, 2004.
- [73] Lamma, Evelina, Paola Mello, and Fabrizio Riguzzi. "A System for Measuring Function Points from an ER-DFD Specification." *The Computer Journal* 47.3 (2004): 358-372.
- [74] Umbers, Paul, and Gail Miles. "Resource estimation for Web applications." *Software Metrics, 2004. Proceedings. 10th International Symposium on*. IEEE, 2004.
- [75] Sanchez-Segura, Maria I., et al. "Virtual reality systems estimation vs. traditional systems estimation." *Journal of Systems and Software* 72.2 (2004): 187-194.

- [76] Al-Hajri, Mohammed Abdullah, et al. "Modification of standard function point complexity weights system." *Journal of Systems and Software* 74.2 (2005): 195-206.
- [77] Diab, H., et al. " μ cROSE: automated measurement of COSMIC-FFP for Rational Rose RealTime." *Information and Software Technology* 47.3 (2005): 151-166.
- [78] Harput, Vahan, Hermann Kaindl, and Stefan Kramer. "Extending function point analysis to object-oriented requirements specifications." *Software Metrics*, 2005. 11th IEEE International Symposium. IEEE, 2005.
- [79] Kralj, Tomaz, et al. "Improved standard FPA method-resolving problems with upper boundaries in the rating complexity process." *Journal of Systems and Software* 77.2 (2005): 81-90.
- [80] Santillo, Luca, Massimiliano Conte, and Roberto Meli. "Early& Quick function point: sizing more with less." *Software Metrics*, 2005. 11th IEEE International Symposium. IEEE, 2005.
- [81] Sheetz, Steven, Linda Wallace, and David Henderson. "Understanding Manager and Developer Perceptions of the Relative Advantage, Compatibility, and Complexity of Function Points and Source Lines of Code." *AMCIS 2005 Proceedings* (2005): 384.
- [82] Zivkovic, Ales, Ivan Rozman, and Marjan Hericko. "Automated software size estimation based on function points using UML models." *Information and Software Technology* 47.13 (2005): 881-890.
- [83] Hericko, Marjan, Ivan Rozman, and Ales Zivkovic. "A formal representation of functional size measurement methods." *Journal of Systems and Software* 79.9 (2006): 1341-1358.
- [84] Abrahao, Silvia, and Geert Poels. "Experimental evaluation of an object-oriented function point measurement procedure." *Information and Software Technology* 49.4 (2007): 366-380.
- [85] Gencel, Cigdem, and Onur Demirors. "Conceptual differences among functional size measurement methods." *null* (2007): 305-313.
- [86] Del Bianco, Vieri, Claudio Gentile, and Luigi Lavazza. "An Evaluation of Function Point Counting Based on Measurement-Oriented Models." *EASE*. 2008.
- [87] Cuadrado-Gallego, Juan J., et al. "IFPUG-cosmic statistical conversion." *Software Engineering and Advanced Applications*, 2008. SEAA'08. 34th Euromicro Conference. IEEE, 2008.

- [88] Lavazza, Luigi A., Vieri Del Bianco, and Carla Garavaglia. "Model-based functional size measurement." Proceedings of the Second ACM-IEEE international symposium on Empirical software engineering and measurement. ACM, 2008.
- [89] Abrahao, Silvia, and Emilio Insfran. "A metamodeling approach to estimate software size from requirements specifications." 34th Euromicro Conference Software Engineering and Advanced Applications. IEEE, 2008.
- [90] Xia, Wei, et al. "A new calibration for Function Point complexity weights." Information and Software Technology 50.7-8 (2008): 670-683.
- [91] Park, Heejun, and Seung Baek. "An empirical validation of a neural network model for software effort estimation." Expert Systems with Applications 35.3 (2008): 929-937.
- [92] Cuadrado-Gallego, Juan J., Fernando Machado-Piriz, and Javier Aroba-Paez. "On the conversion between IFPUG and COSMIC software functional size units: A theoretical and empirical study." Journal of Systems and Software 81.5 (2008): 661-672.
- [93] Hericko, Marjan, and Ales Zivkovic. "The size and effort estimates in iterative development." Information and Software Technology 50.7-8 (2008): 772-781.
- [94] Kampstra, P., and C. Verhoef. "Reliability of function point counts." Communications of the ACM, Amsterdam (2009): 85-97.
- [95] Lavazza, Luigi, and Carla Garavaglia. "Using function points to measure and estimate real-time and embedded software: Experiences and guidelines." Empirical Software Engineering and Measurement, 2009. ESEM 2009. 3rd International Symposium on. IEEE, 2009.
- [96] Demirors, Onur, and Cigdem Gencel. "Conceptual association of functional size measurement methods." IEEE software 26.3 (2009).
- [97] Del Bianco, Vieri, and Luigi Lavazza. "Applying the COSMIC functional size measurement method to problem frames." Engineering of Complex Computer Systems, 2009 14th IEEE International Conference on. IEEE, 2009.
- [98] Bagheri, Saeed, Krishna Ratakonda, and Rakesh Mohan. "Variable Productivity Adjustment Estimation for Function Point Project Delivery." Data Mining Workshops (ICDMW), 2010 IEEE International Conference on. IEEE, 2010.

- [99] Lind, Kenneth, and Rogardt Heldal. "Categorization of real-time software components for code size estimation." Proceedings of the 2010 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement. ACM, 2010.
- [100] Lavazza, Luigi, and Gabriela Robiolo. "Introducing the evaluation of complexity in functional size measurement: a UML-based approach." Proceedings of the 2010 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement. ACM, 2010.
- [101] Abrahao, Silvia, Jaime Gomez, and Emilio Insfran. "Validating a size measure for effort estimation in model-driven Web development." Information Sciences 180.20 (2010): 3932-3954.
- [102] Edagawa, T., et al. "Function point measurement from Web application source code based on screen transitions and database accesses." Journal of Systems and Software 84.6 (2011): 976-984.
- [103] El-Sebakhy, Emad A. "Functional networks as a novel data mining paradigm in forecasting software development efforts." Expert Systems with Applications 38.3 (2011): 2187-2194.
- [104] Lind, Kenneth, and Rogardt Heldal. "A model-based and automated approach to size estimation of embedded software components." International Conference on Model Driven Engineering Languages and Systems. Springer, Berlin, Heidelberg, 2011.
- [105] Wilkie, F. George, et al. "The value of software sizing." Information and Software Technology 53.11 (2011): 1236-1249.
- [106] Choi, Soonhwang, Sooyong Park, and Vijayan Sugumaran. "A rule-based approach for estimating software development cost using function point and goal and scenario based requirements." Expert Systems with Applications 39.1 (2012): 406-418.
- [107] Abualkishik, Abedallah Zaid, et al. "An exploratory study on the accuracy of FPA to COSMIC measurement method conversion types." Information and Software Technology 54.11 (2012): 1250-1264.
- [108] Tsunoda, Masateru, et al. "Revisiting software development effort estimation based on early phase development activities." 2013 10th Working Conference on Mining Software Repositories (MSR). IEEE, 2013.

- [109] Staples, Mark, et al. "Formal specifications better than function points for code sizing." Proceedings of the 2013 International Conference on Software Engineering. IEEE Press, 2013.
- [110] Pai, Dinesh R., Kevin S. McFall, and Girish H. Subramanian. "Software effort estimation using a neural network ensemble." Journal of Computer Information Systems 53.4 (2013): 49-58.
- [111] Lavazza, Luigi, Sandro Morasca, and Gabriela Robiolo. "Towards a simplified definition of Function Points." Information and Software Technology 55.10 (2013): 1796-1809.
- [112] Quesada-Lopez, Christian, and Marcelo Jenkins. "Function point structure and applicability validation using the ISBSG dataset: a replicated study." Proceedings of the 8th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement. ACM, 2014.
- [113] Del Bianco, Vieri, et al. "Model-based early and rapid estimation of COSMIC functional size-An experimental evaluation." Information and Software Technology 56.10 (2014): 1253-1267.
- [114] Paz, Freddy, Claudia Zapata, and Jose Antonio Pow-Sang. "An approach for effort estimation in incremental software development using cosmic function points." Proceedings of the 8th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement. ACM, 2014.
- [115] Ceke, Denis, and Boris Milasinovic. "Early effort estimation in web application development." Journal of Systems and Software 103 (2015): 219-237.
- [116] de Freitas Junior, Marcos, Marcelo Fantinato, and Violeta Sun. "Improvements to the function point analysis method: A systematic literature review." IEEE Transactions on Engineering Management 62.4 (2015): 495-506.
- [117] Huijgens, Hennie, Georgios Gousios, and Arie van Deursen. "Pricing via functional size-a case study of a company's portfolio of 77 outsourced projects." Empirical Software Engineering and Measurement (ESEM), 2015 ACM/IEEE International Symposium on. IEEE, 2015.
- [118] Sellami, Asma, et al. "A measurement method for sizing the structure of UML sequence diagrams." Information and Software Technology 59 (2015): 222-232.

- [119] Hira, Anandi, and Barry Boehm. "Function point analysis for software maintenance." Proceedings of the 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement. ACM, 2016.
- [120] Hira, Anandi, and Barry Boehm. "Using Software Non-Functional Assessment Process to Complement Function Points for Software Maintenance." Proceedings of the 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement. ACM, 2016.
- [121] Bagriyanik, Selami, and Adem Karahoca. "Automated COSMIC Function Point measurement using a requirements engineering ontology." *Information and Software Technology* 72 (2016): 189-203.
- [122] Di Martino, Sergio, et al. "Web effort estimation: function point analysis vs. COSMIC." *Information and Software Technology* 72 (2016): 90-109.
- [123] Huijgens, Hennie, et al. "An exploratory study on functional size measurement based on code." Proceedings of the International Conference on Software and Systems Process. ACM, 2016.
- [124] Ochodek, Mirosaw. "Functional size approximation based on use-case names." *Information and Software Technology* 80 (2016): 73-88.
- [125] Abrahao, Silvia, et al. "Definition and evaluation of a COSMIC measurement procedure for sizing Web applications in a model-driven development environment." *Information and Software Technology* 104 (2018): 144-161.
- [126] Abualkishik, Abedallah Zaid, and Luigi Lavazza. "IFPUG Function Points to COSMIC Function Points convertibility: A fine-grained statistical approach." *Information and Software Technology* 97 (2018): 179-191.

付録

- [I] Albrecht, Allan J., and John E. Gaffney. "Software function, source lines of code, and development effort prediction: a software science validation." *IEEE transactions on software engineering* 6 (1983): 639-648.
- [II] Behrens, Charles A. "Measuring the productivity of computer systems development activities with function points." *IEEE Transactions on Software Engineering* 6 (1983): 648-652.
- [III] Kemerer, Chris F. "An empirical validation of software cost estimation models." *Communications of the ACM* 30.5 (1987): 416-429.
- [IV] Symons, Charles R. "Function point analysis: difficulties and improvements." *IEEE transactions on software engineering* 14.1 (1988): 2-11.
- [V] Verner, June M., et al. "Technology dependence in function point analysis: a case study and critical review." *Software Engineering, 1989. 11th International Conference on. IEEE, 1989.*
- [VI] Low, Graham C., and D. Ross Jeffery. "Function points in the estimation and evaluation of the software process." *IEEE Transactions on Software Engineering* 1 (1990): 64-71.
- [VII] Reifer, Donald J. "Asset-R: A function point sizing tool for scientific and real-time systems." *Journal of Systems and Software* 11.3 (1990): 159-171.
- [VIII] Heemstra, F. J., and R. J. Kusters. "Function point analysis: Evaluation of a software cost estimation model." *European Journal of Information Systems* 1.4 (1991): 229-237.
- [IX] Guerra, L. G., and H. S. Lobos. "An Empirical Validation of a Software Cost Model." *IFAC Proceedings Volumes* 23.5 (1990): 57-62.
- [X] Bock, Douglas B., and Robert Klepper. "FP-S: A simplified function point counting method." *Journal of Systems and Software* 18.3 (1992): 245-254.
- [XI] Kemerer, Chris F., and Benjamin S. Porter. "Improving the reliability of function point measurement: an empirical study." *IEEE Transactions on Software Engineering* 18.11 (1992): 1011-1024.
- [XII] Rask, Raimo, Petteri Laamanen, and Kalle Lyytinen. "A comparison of Albrecht's function point and Symons' Mark II metrics." *ICIS. 1992.*

- [XIII] Betteridge, R. "Successful experience of using function points to estimate project costs early in the life-cycle." *Information and Software Technology* 34.10 (1992): 655-658.
- [XIV] Kitchenham, Barbara, and Kari Kansala. "Inter-item correlations among function points." *Software Metrics Symposium, 1993. Proceedings., First International.* IEEE, 1993.
- [XV] Jeffery, D. Ross, Graham C. Low, and M. Barnes. "A comparison of function point counting techniques." *IEEE Transactions on Software Engineering* 19.5 (1993): 529-532.
- [XVI] Rask, Raimo, Petteri Laamanen, and K. Lyytinen. "Simulation and comparison of Albrecht's function point and DeMarco's function bang metrics in a CASE environment." *IEEE Transactions on Software Engineering* 19.7 (1993): 661-671.
- [XVII] Hughes, R. "An exploration of the problems of using function points mark II." *Information Systems Journal* 4.3 (1994): 169-183.
- [XVIII] Matson, Jack E., Bruce E. Barrett, and Joseph M. Mellichamp. "Software development cost estimation using function points." *IEEE Transactions on Software Engineering* 20.4 (1994): 275-287.
- [XIX] Abran, Alain, and Pierre N. Robillard. "Function points: a study of their measurement processes and scale transformations." *Journal of Systems and Software* 25.2 (1994): 171-184.
- [XX] Kitchenham, Barbara, Shari Lawrence Pfleeger, and Norman Fenton. "Towards a framework for software measurement validation." *IEEE Transactions on software Engineering* 21.12 (1995): 929-944.
- [XXI] Jones, Capers. "Backfiring: Converting lines of code to function points." *Computer* 28.11 (1995): 87-88.
- [XXII] Jones, Capers. "Determining software schedules." *Computer* 28.2 (1995): 73-75.
- [XXIII] Abran, Alain, and Pierre N. Robillard. "Function points analysis: An empirical study of its measurement processes." *IEEE Transactions on Software Engineering* 22.12 (1996): 895-910.
- [XXIV] Subramanian, Girish H., and George E. Zarnich. "An examination of some software development effort and productivity determinants in ICASE tool projects." *Journal of Management Information Systems* 12.4 (1996): 143-160.

- [XXV] Finnie, Gavin R., Gerhard E. Wittig, and Jean-Marc Desharnais. "A comparison of software effort estimation techniques: using function points with neural networks, case-based reasoning and regression models." *Journal of systems and software* 39.3 (1997): 281-289.
- [XXVI] Furey, Sean. "Why we should use function points [software metrics]." *IEEE Software* 14.2 (1997): 28.
- [XXVII] Shoval, Peretz, and Opher 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* 39 (1997): 855-865.
- [XXVIII] Dolado, Jose Javier. "A study of the relationships among Albrecht and Mark II function points, lines of code 4GL and effort." *Journal of Systems and Software* 37.2 (1997): 161-173.
- [XXIX] Antoniol, Giuliano, et al. "Adapting function points to object oriented information systems." *International Conference on Advanced Information Systems Engineering*. Springer, Berlin, Heidelberg, 1998.
- [XXX] Fujiwara, Fumiko, Takushi Goto, and Sadao Araki. "Examples of applying software estimate tool." *Software Engineering, 1998. Proceedings of the 1998 International Conference on*. IEEE, 1998.
- [XXXI] Fleck, Robert A., and Jr. "Managing programmer resources in a maintenance environment with function points." *Industrial Management & Data Systems* 98.2 (1998): 63-70.
- [XXXII] Yau, Chuk, and Ho-Leung Tsoi. "Modelling the probabilistic behaviour of function point analysis." *Information and Software Technology* 40.2 (1998): 59-68.
- [XXXIII] Horgan, Gerard, Souheil Khaddaj, and Peter Forte. "Construction of an FPA-type metric for early lifecycle estimation." *Information and Software Technology* 40.8 (1998): 409-415.
- [XXXIV] Bok, Hai Suan, and Krishnamurthy S. Raman. "Software engineering productivity measurement using function points: a case study." *Journal of Information Technology* 15.1 (2000): 79-90.
- [XXXV] Lokan, Christopher J. "An empirical analysis of function point adjustment factors." *Information and Software Technology* 42.9 (2000): 649-659.

- [XXXVI] Orr, George, and Thomas E. Reeves. "Function point counting: one program's experience." *Journal of Systems and Software* 53.3 (2000): 239-244.
- [XXXVII] Van Genuchten, Michiel, and Hans Koolen. "On the use of software cost models." *Information & Management* 21.1 (1991): 37-44.
- [XXXVIII] Kusumoto, Shinji, et al. "Function point measurement from Java programs." *Proceedings of the 24th International Conference on Software Engineering*. ACM, 2002.
- [XXXIX] Kitchenham, Barbara, et al. "An empirical study of maintenance and development estimation accuracy." *Journal of systems and software* 64.1 (2002): 57-77.
- [XL] Calazans, Angelica Toffano, Kathia Marcal de Oliveira, and Rildo Ribeiro dos Santos. "Adapting function point analysis to estimate data mart size." *Software Metrics, 2004. Proceedings. 10th International Symposium on*. IEEE, 2004.
- [XLI] Cantone, Giovanni, Davide Pace, and Giuseppe Calavaro. "Applying function point to unified modeling language: Conversion model and pilot study." *Software Metrics, 2004. Proceedings. 10th International Symposium on*. IEEE, 2004.
- [XLII] Lamma, Evelina, Paola Mello, and Fabrizio Riguzzi. "A System for Measuring Function Points from an ER-DFD Specification." *The Computer Journal* 47.3 (2004): 358-372.
- [XLIII] Umbers, Paul, and Gail Miles. "Resource estimation for Web applications." *Software Metrics, 2004. Proceedings. 10th International Symposium on*. IEEE, 2004.
- [XLIV] Sanchez-Segura, Maria I., et al. "Virtual reality systems estimation vs. traditional systems estimation." *Journal of Systems and Software* 72.2 (2004): 187-194.
- [XLV] Al-Hajri, Mohammed Abdullah, et al. "Modification of standard function point complexity weights system." *Journal of Systems and Software* 74.2 (2005): 195-206.
- [XLVI] Diab, H., et al. " μ cROSE: automated measurement of COSMIC-FFP for Rational Rose RealTime." *Information and Software Technology* 47.3 (2005): 151-166.
- [XLVII] Harput, Vahan, Hermann Kaindl, and Stefan Kramer. "Extending function point analysis to object-oriented requirements specifications." *Software Metrics, 2005. 11th IEEE International Symposium*. IEEE, 2005.
- [XLVIII] Kralj, Tomaz, et al. "Improved standard FPA method-resolving problems with upper boundaries in the rating complexity process." *Journal of Systems and Software* 77.2 (2005): 81-90.

- [XLIX] Santillo, Luca, Massimiliano Conte, and Roberto Meli. "Early& Quick function point: sizing more with less." *Software Metrics*, 2005. 11th IEEE International Symposium. IEEE, 2005.
- [L] Sheetz, Steven, Linda Wallace, and David Henderson. "Understanding Manager and Developer Perceptions of the Relative Advantage, Compatibility, and Complexity of Function Points and Source Lines of Code." *AMCIS 2005 Proceedings* (2005): 384.
- [LI] Zivkovic, Ales, Ivan Rozman, and Marjan Hericko. "Automated software size estimation based on function points using UML models." *Information and Software Technology* 47.13 (2005): 881-890.
- [LII] Hericko, Marjan, Ivan Rozman, and Ales Zivkovic. "A formal representation of functional size measurement methods." *Journal of Systems and Software* 79.9 (2006): 1341-1358.
- [LIII] Abrahao, Silvia, and Geert Poels. "Experimental evaluation of an object-oriented function point measurement procedure." *Information and Software Technology* 49.4 (2007): 366-380.
- [LIV] Gencel, Cigdem, and Onur Demirors. "Conceptual differences among functional size measurement methods." *null* (2007): 305-313.
- [LV] Del Bianco, Vieri, Claudio Gentile, and Luigi Lavazza. "An Evaluation of Function Point Counting Based on Measurement-Oriented Models." *EASE*. 2008.
- [LVI] Cuadrado-Gallego, Juan J., et al. "IFPUG-cosmic statistical conversion." *Software Engineering and Advanced Applications*, 2008. SEAA'08. 34th Euromicro Conference. IEEE, 2008.
- [LVII] Lavazza, Luigi A., Vieri Del Bianco, and Carla Garavaglia. "Model-based functional size measurement." *Proceedings of the Second ACM-IEEE international symposium on Empirical software engineering and measurement*. ACM, 2008.
- [LVIII] Abrahao, Silvia, and Emilio Insfran. "A metamodeling approach to estimate software size from requirements specifications." *34th Euromicro Conference Software Engineering and Advanced Applications*. IEEE, 2008.
- [LIX] Xia, Wei, et al. "A new calibration for Function Point complexity weights." *Information and Software Technology* 50.7-8 (2008): 670-683.

- [LX] Park, Heejun, and Seung Baek. "An empirical validation of a neural network model for software effort estimation." *Expert Systems with Applications* 35.3 (2008): 929-937.
- [LXI] Cuadrado-Gallego, Juan J., Fernando Machado-Piriz, and Javier Aroba-Paez. "On the conversion between IFPUG and COSMIC software functional size units: A theoretical and empirical study." *Journal of Systems and Software* 81.5 (2008): 661-672.
- [LXII] Hericko, Marjan, and Ales Zivkovic. "The size and effort estimates in iterative development." *Information and Software Technology* 50.7-8 (2008): 772-781.
- [LXIII] Kampstra, P., and C. Verhoef. "Reliability of function point counts." *Communications of the ACM, Amsterdam* (2009): 85-97.
- [LXIV] Lavazza, Luigi, and Carla Garavaglia. "Using function points to measure and estimate real-time and embedded software: Experiences and guidelines." *Empirical Software Engineering and Measurement, 2009. ESEM 2009. 3rd International Symposium on. IEEE, 2009.*
- [LXV] Demirors, Onur, and Cigdem Gencel. "Conceptual association of functional size measurement methods." *IEEE software* 26.3 (2009).
- [LXVI] Del Bianco, Vieri, and Luigi Lavazza. "Applying the COSMIC functional size measurement method to problem frames." *Engineering of Complex Computer Systems, 2009 14th IEEE International Conference on. IEEE, 2009.*
- [LXVII] Bagheri, Saeed, Krishna Ratakonda, and Rakesh Mohan. "Variable Productivity Adjustment Estimation for Function Point Project Delivery." *Data Mining Workshops (ICDMW), 2010 IEEE International Conference on. IEEE, 2010.*
- [LXVIII] Lind, Kenneth, and Rogardt Haldal. "Categorization of real-time software components for code size estimation." *Proceedings of the 2010 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement. ACM, 2010.*
- [LXIX] Lavazza, Luigi, and Gabriela Robiolo. "Introducing the evaluation of complexity in functional size measurement: a UML-based approach." *Proceedings of the 2010 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement. ACM, 2010.*
- [LXX] Abrahao, Silvia, Jaime Gomez, and Emilio Insfran. "Validating a size measure for effort estimation in model-driven Web development." *Information Sciences* 180.20 (2010): 3932-3954.

- [LXXI] Edagawa, T., et al. "Function point measurement from Web application source code based on screen transitions and database accesses." *Journal of Systems and Software* 84.6 (2011): 976-984.
- [LXXII] El-Sebakhy, Emad A. "Functional networks as a novel data mining paradigm in forecasting software development efforts." *Expert Systems with Applications* 38.3 (2011): 2187-2194.
- [LXXIII] Lind, Kenneth, and Rogardt Heldal. "A model-based and automated approach to size estimation of embedded software components." *International Conference on Model Driven Engineering Languages and Systems*. Springer, Berlin, Heidelberg, 2011.
- [LXXIV] Wilkie, F. George, et al. "The value of software sizing." *Information and Software Technology* 53.11 (2011): 1236-1249.
- [LXXV] Choi, Soonhwang, Sooyong Park, and Vijayan Sugumaran. "A rule-based approach for estimating software development cost using function point and goal and scenario based requirements." *Expert Systems with Applications* 39.1 (2012): 406-418.
- [LXXVI] Abualkishik, Abedallah Zaid, et al. "An exploratory study on the accuracy of FPA to COSMIC measurement method conversion types." *Information and Software Technology* 54.11 (2012): 1250-1264.
- [LXXVII] Tsunoda, Masateru, et al. "Revisiting software development effort estimation based on early phase development activities." *2013 10th Working Conference on Mining Software Repositories (MSR)*. IEEE, 2013.
- [LXXVIII] Staples, Mark, et al. "Formal specifications better than function points for code sizing." *Proceedings of the 2013 International Conference on Software Engineering*. IEEE Press, 2013.
- [LXXIX] Pai, Dinesh R., Kevin S. McFall, and Girish H. Subramanian. "Software effort estimation using a neural network ensemble." *Journal of Computer Information Systems* 53.4 (2013): 49-58.
- [LXXX] Lavazza, Luigi, Sandro Morasca, and Gabriela Robiolo. "Towards a simplified definition of Function Points." *Information and Software Technology* 55.10 (2013): 1796-1809.
- [LXXXI] Quesada-Lopez, Christian, and Marcelo Jenkins. "Function point structure and applicability validation using the ISBSG dataset: a replicated study." *Proceedings of the 8th ACM/IEEE*

- International Symposium on Empirical Software Engineering and Measurement. ACM, 2014.
- [LXXXII] Del Bianco, Vieri, et al. "Model-based early and rapid estimation of COSMIC functional size-An experimental evaluation." *Information and Software Technology* 56.10 (2014): 1253-1267.
- [LXXXIII] Paz, Freddy, Claudia Zapata, and Jose Antonio Pow-Sang. "An approach for effort estimation in incremental software development using cosmic function points." *Proceedings of the 8th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*. ACM, 2014.
- [LXXXIV] Ceke, Denis, and Boris Milasinovic. "Early effort estimation in web application development." *Journal of Systems and Software* 103 (2015): 219-237.
- [LXXXV] de Freitas Junior, Marcos, Marcelo Fantinato, and Violeta Sun. "Improvements to the function point analysis method: A systematic literature review." *IEEE Transactions on Engineering Management* 62.4 (2015): 495-506.
- [LXXXVI] Huijgens, Hennie, Georgios Gousios, and Arie van Deursen. "Pricing via functional size-a case study of a company's portfolio of 77 outsourced projects." *Empirical Software Engineering and Measurement (ESEM), 2015 ACM/IEEE International Symposium on*. IEEE, 2015.
- [LXXXVII] Sellami, Asma, et al. "A measurement method for sizing the structure of UML sequence diagrams." *Information and Software Technology* 59 (2015): 222-232.
- [LXXXVIII] Hira, Anandi, and Barry Boehm. "Function point analysis for software maintenance." *Proceedings of the 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*. ACM, 2016.
- [LXXXIX] Hira, Anandi, and Barry Boehm. "Using Software Non-Functional Assessment Process to Complement Function Points for Software Maintenance." *Proceedings of the 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*. ACM, 2016.
- [XC] Bagriyanik, Selami, and Adem Karahoca. "Automated COSMIC Function Point measurement using a requirements engineering ontology." *Information and Software Technology* 72 (2016): 189-203.

- [XCI] Di Martino, Sergio, et al. "Web effort estimation: function point analysis vs. COSMIC." *Information and Software Technology* 72 (2016): 90-109.
- [XCII] Huijgens, Hennie, et al. "An exploratory study on functional size measurement based on code." *Proceedings of the International Conference on Software and Systems Process*. ACM, 2016.
- [XCIII] Ochodek, Mirosaw. "Functional size approximation based on use-case names." *Information and Software Technology* 80 (2016): 73-88.
- [XCIV] Abrahao, Silvia, et al. "Definition and evaluation of a COSMIC measurement procedure for sizing Web applications in a model-driven development environment." *Information and Software Technology* 104 (2018): 144-161.
- [XCV] Abualkishik, Abedallah Zaid, and Luigi Lavazza. "IFPUG Function Points to COSMIC Function Points convertibility: A fine-grained statistical approach." *Information and Software Technology* 97 (2018): 179-191.