

FP 計測手法における FP 規模と工数の相関の差

東海大学理学部
IPA/SEC 専門委員

古山 恒夫



IPA/SEC で収集したエンタプライズ系ソフトウェアプロジェクトデータから、FP 計測手法ごとに FP 規模と実績工数の相関分析を行った。新規開発プロジェクトでは IFPUG 法の相関が最も高く、簡易な FP 計測手法である SPR 法や NESMA 概算法よりも決定係数でそれぞれ 0.169 と 0.250 大きな値となった。同じ IFPUG 法でも製造業のプロジェクト群の方が金融・保険業よりも相関が高い傾向が見られた。

1. はじめに

FP 計測手法として、IFPUG 法、COSMIC 法、フルファンクションポイント法、フィーチャポイント法、Mark II 法、NESMA 概算法、SPR 法などさまざまな分野のソフトウェアに対してさまざまな計測手法が提案されている [富野 2010]。このうち、エンタプライズ系ソフトウェアに対しては、主に IFPUG (International Function Point User Group) が提案した IFPUG 法が使われているようである。IFPUG 法は FP のカウント方法が詳細に定義されているため、FP 計測の精度は向上するが、逆に設計仕様書が明確にならないと正確な計測ができないという問題がある。また FP 計測の工数もかかる。これらの問題に対処するために、計算を簡略化することにより、開発の早期に比較的少ない工数で FP を求めることができるようにした NESMA 概算法や SPR 法も使われている。ちなみに IPA/SEC で収集したエンタプライズ系ソフトウェアプロジェクトデータにおける FP 計測手法の分布を調べてみると、IFPUG 法が 32%、SPR 法が 21%、NESMA 概算法が 10%、その他の FP 計測手法が 33% となっている (FP 計測手法の記載のないプロジェクトが 4% ある)。

本稿ではエンタプライズ系ソフトウェアプロジェクトで使われている、これら 4 つの FP 計測手法における FP 規模と工数の相関分析を行った結果を報告する。この相関係数の値が大きいということは、FP 計測手法による工数予測の精度が高いということになる。その差を比較することにより、早期見積りにおける見積もり損失の大き

かな目安を得ることができると考えられる。

2 章で 4 つの FP 計測手法の概要を、3 章で分析結果を示す。3 章では各 FP 計測手法の適用プロジェクトの大きさの比較、相関分析の結果、より詳細な層別による分析結果を示す。4 章で分析結果に対する考察を、5 章でまとめを述べる。

2. FP 計測手法の概要

2.1 IFPUG 法

IFPUG 法では次の手順で機能量 (FP) を計算する [JISX0142], [富野 2010]。

- ① 扱うデータを外部入力 (EI)、外部出力 (EO)、外部照会 (EQ)、内部論理ファイル (ILF)、外部インタフェースファイル (EIF) の 5 つのファンクション型 (以下ではタイプと略称) に分ける。
- ② 扱うデータごとに、「データ項目数」とそのデータに関連する「レコード種類数」を求め、それによってそのデータのファンクションの複雑さを 3 段階 (低、中、高) に分ける。3 段階の分け方は、同じ「データ項目数」及び「レコード種類数」であってもデータのタイプによって異なる。
- ③ 各データに 3 段階のファンクションの複雑さに応じた未調整 FP (重み係数) を掛けて合計し、システム全体の未調整 FP を求める。ファンクションの複雑さのレベルが同じでも、データのタイプによって未調整 FP の値は異なる。
- ④ ①～③までの計算と独立に、対象とするシステムの特徴を 14 の観点から 0～5 の 6 段階評価し、合計

する。

⑤ システム特性係数 = 0.65 + ④の合計値 × 0.01 を計算する。

⑥ FP = システム特性係数 × 未調整 FP

JIS 規格として制定されているのは①から③までである。

2.2 NESMA 概算法

NESMA 概算法はオランダの NESMA (Netherlands Software Metrics Association) が提案した手法であり、開発初期段階でも利用できる [NESMA]。NESMA 概算法では、IFPUG 法における②と③を一本化し、各タイプのデータに与える複雑さは3段階ではなく、データのタイプごとに固定したものとする。すなわち、外部入力、外部出力、外部照会の3つのタイプのデータに対しては、常に「中」の複雑さを与え、内部論理ファイル、外部インターフェースファイルの2つタイプのデータに対しては常に「低」の複雑さを与える。これによって、扱うデータの詳細がわからない段階でも未調整 FP を計算することができるようになる。

2.3 SPR 法

SPR 法は SPR 社がファンクションポイントの計算工数削減をするために計算を単純化したものである [富野 2010]。SPR 法では IFPUG 法における②と③を一本化し、データのタイプによらず複雑さは常に「中」とする。

また、システム特性に基づく調整係数の算出は、IFPUG 法の④とは異なり、問題の複雑さ、データの複雑さの2つの視点から5段階評価し、それらの合計値に応じた調整係数を求める。

2.4 「その他」の FP 計測手法

エンタプライズ系のプロジェクトデータには、FP 計測手法の選択肢として「その他」というものがある。「その他」は主に IFPUG 法をベースに各社が独自に開発した FP 計測手法である。これを本文中では「その他」とカギ括弧付きで表す。

3. 分析対象プロジェクト

3.1 分析対象プロジェクト

分析は、IPA/SEC で収集したエンタプライズ系ソフトウェアプロジェクト 3,089 件のデータ [IPA2012] のうち、次の条件を満たすプロジェクトを対象として分析を行った。

① 開発 5 工程が揃っている。

② FP の実測値 (5001_FP 実測値_調整前) がある。

③ FP 計測手法が明記されている。

④ 層のデータ数が 30 件以上である。

なお、工数のデータはすべてのプロジェクトから報告されている。

3.2 層別プロジェクト件数

開発種別及び FP 計測種別ごとのプロジェクト件数を表 1 に示す。各層とも散布図の目視により若干の外れ値を除いている。なお、以下の表中では、新規開発を「新規」、改修・保守を「改修」と略記する。

改修・保守の NESMA 概算のデータ数が 10 であることを除いて、新規開発及び改修・保守ではすべての層でデータ数が 30 以上であった。したがって、分析はこの 7 つの層を対象として行う。再開発及び拡張を FP 計測種別で層別すると各層のデータ数はすべて 30 件未満となったので、今回は分析対象とはしていない。

表 1 層別プロジェクト件数

開発種別	FP 計測種別				合計
	IFPUG	SPR	NESMA	その他	
新規	200 (4)	77 (2)	68	169 (1)	514 (7)
改修	52 (1)	34	10	75	171 (1)

(注) () 内は外れ値として除いたもの (別掲)

3.3 変数

FP 規模、工数はそれぞれ次のものを用いる。

- ・ FP 規模 (実績値) : 5001_FP 実測値_調整前
- ・ 工数 (実績値) : 開発 5 工程工数

ただし、分析に際しては上記のデータをすべて常用対数で対数変換したものを用いる。

4. 分析結果

4.1 有意水準

一般に、P 値が 5% あるいは 1% 以下である場合に「(有意水準 5% あるいは有意水準 1% で) 統計的に有意である」という。しかし、プロジェクトデータを分析した結果、P 値が 5% 以下になるようなケースはそれほど多くはない。また、P 値が 5% を超えるような結果であっても、ソフトウェアプロジェクト管理では有用と思われる情報もあると考えられる。そこで、本稿では、P 値が 20% 以下の場合、すなわち 5 回のうち 1 回は誤った情報を与える可能性がある場合も、分析結果は有用な情報であると考えて、「・・・の傾向がある」という表現で表すことにする。

4.2 適用プロジェクトの大きさの比較

3.2で述べた7つの層のプロジェクトの大きさを比較するために、FP規模・工数の各ランクをまとめたものを表2に示す。「ランク」で同じ記号のものは、t検定(有意水準5%)の結果各層の平均値同士で有意差がないことを表している。例えば、表2の「新規・IFPUG法」と「新規・NESMA概算法」は同じ「A」であるので、「新規・IFPUG法」の平均値(常用対数変換後の値で3.048であった)が「新規・NESMA概算法」の平均値(2.964であった)よりも大きいけれど各母集団のFP規模の平均値に統計的に有意な差があるとは認められない、ということを表している。

表2 プロジェクト規模の比較

開発種別	FP計測手法	データ数	ランク	
			FP規模	工数
新規	IFPUG	200	A	A
新規	NESMA	68	A	A
改修	IFPUG	52	B	B
新規	その他	169	B	C
新規	SPR	77	B	D
改修	その他	75	C	D
改修	SPR	34	C	E

表2から次のことがわかる。

- (1) FP規模と工数の大きさの順序は変わらない。
- (2) 新規開発・IFPUG法と新規開発・NESMA概算法が最も大きなプロジェクト群に適用されている。2つのグループの間にはFP規模・工数のいずれにも差が見られない。
- (3) 改修・保守・IFPUG法、新規開発・「その他」、新規開発・SPR法の順で適用プロジェクト群の大きさが小さくなる。
- (4) 改修・保守の「その他」と改修・保守のSPR法は比較的小さいプロジェクトに適用されている。

4.3 相関分析

各層のFP規模と工数の相関係数を表3に示す。相関係数をz変換すると正規分布に近似できるので、z変換後の値に対して相互にt検定を行って有意差の有無を調べ、ランクを求めた。ランクは新規開発・IFPUG法だけがAで他の層はすべて同一のBである。すなわち、新規開発・IFPUG法の相関だけが他の層と統計的に有意なほど高く、それ以外の層同士では統計的に有意な差があるとは言えないことを示している。従って、例えば改修・保守・IFPUG法の相関係数0.807と新規開発・NESMA概算法の相関係数0.735の間に統計的な有意差はない。

表3 相関係数と決定係数

開発種別	FP計測手法	データ数	ランク	相関係数	決定係数(*1)
新規	IFPUG	200	A	0.886	0.784
改修	IFPUG	52	B	0.807	0.645
新規	その他	169	B	0.800	0.638
新規	SPR	77	B	0.788	0.615
改修	その他	75	B	0.757	0.568
改修	SPR	34	B	0.752	0.553
新規	NESMA	68	B	0.735	0.534

(*1) 自由度調整済のもの

表3から次のことがわかる。

- (1) 新規開発では、IFPUG法の決定係数(自由度調整済)が0.784に対し、NESMA概算法では0.534、SPR法では0.615となっている。すなわちIFPUG法ではFP規模だけで工数の78.4%が決定されるのに対し、NESMA概算法では53.4%、SPR法では61.5%であり、その差はそれぞれ25%と17%ある。これが早期にFP規模を見積もるために失うものと考えられる。
- (2) 同じIFPUG法でも、改修・保守では、新規開発の場合に比べて決定係数は0.139(14%)小さい。

4.4 業種別相関分析

4.2で述べた結果をより詳細に分析するために、開発5工程のデータを業種別に分けて、FP計測種別ごとのFP規模と工数の相関分析を行った。

FP計測種別ごとに業種別のデータ数を調べたところ30件以上データの揃っている層は「製造業」と「金融・保険業」しかなかった。したがって、業種としてはこの2つのみを分析対象とした。

新規開発の業種別の相関係数及び相関係数間の差の有無のP値を表4に示す。表4の「全業種」に示されている各FP計測手法の値は表3からの再掲である。

表4から同一のFP計測手法でも業種間で相関係数に差があることがわかる。すなわち、IFPUG法では、製造業での相関係数が金融・保険業での相関係数よりも大きい傾向にあり(t検定のP値0.168)、「その他」では製造業の相関係数は金融・保険業の相関係数よりも大きい(t検定のP値0.001)。

IFPUG法と「その他」の相関係数は同一業種内でみても差がある。製造業ではその差のP値は0.159であり、IFPUG法の方が「その他」よりも相関係数が高い傾向があると言える。金融・保険業ではその差のP値は0.001であり、有意水準1%でIFPUG法の方が「その他」よりも相関係数が高いと言える。

表4 業種で層別した相関係数とその差の検定

業種 FP 計測手法	全業種 (参考)	製造業	金融・ 保険業	P 値 (業種間)
IFPUG	0.886 (200)	0.930 (34)	0.873 (56)	0.168
その他	0.800 (169)	0.874 (68)	0.565 (41)	0.001
P 値 (FP 計 測手法間)	0.004	0.159	0.001	

(注) () 内はデータ数

4.5 アーキテクチャ別分析

開発5工程のデータをアーキテクチャ別に分けて、FP計測種別ごとのFP規模と工数の相関分析を行った。

FP計測種別ごとにアーキテクチャ別のデータ数を調べたところ30件以上データの揃っている層は、新規開発では2階層クライアント/サーバ(以下では、2階層C/Sと略記)、3階層クライアント/サーバ(3階層C/Sと略記)、イントラネット/インターネット(ネット系と略記)の3つだけであった。

新規開発のアーキテクチャ別の相関係数及び相関係数間の差の有無のP値を表5に示す。表5の「全アーキテクチャ」に示されている相関係数とデータ数は表3からの再掲である。

表5の相関係数を数値だけで比較すると、IFPUG法では、ネット系>3階層C/Sであり、「その他」ではネット系>3階層C/S>2階層C/Sであるが、t検定の結果からは、同一のFP計測手法ではアーキテクチャ間に差あるとは言えない。

同一アーキテクチャ内でFP計測手法間の差をみると、ネット系におけるIFPUG法の相関係数がSPR法よりも大きい傾向にある(P値は0.091)ことがわかる。

表5 アーキテクチャ別で層別した相関係数とその差の検定

アーキテクチャ FP 計測手法	全アーキ テクチャ (参考)	2階層C/S (A)	3階層C/S (B)	ネット系 (C)	P 値 (アーキ テクチャ間)
IFPUG (a)	0.886 (200)	-	0.859 (37)	0.905 (128)	>0.2
SPR (b)	0.788 (77)	-	-	0.824 (36)	-
その他 (c)	0.800 (169)	0.797 (55)	0.827 (70)	0.873 (32)	>0.2 (*1)
P 値 (FP 計測 手法間)	(a)と(b) は 0.013 (b)と(c) は >0.2	-	>0.2	(a)と(b) は 0.0091 他の組合 せは >0.2	

(注) () 内はデータ数
(*1) (A), (B), (C) の任意の2つの組合せに対する結果がすべて>0.2である。

5. 考察

(1) FP計測手法による相関係数の違いについて

NESMA概算法やSPR法ではIFPUG法で行う各データ種別の複雑さの評価を省略し、評価をデータ種別ごとに固定している(図1)。それにより、早期にFPを求めることができるとともに、計測工数を削減することもできる。しかし、FPの精度が低下するというデメリットもある。

今回の分析では、新規開発におけるNESMA概算法のFP規模の工数に対する決定係数が新規開発のIFPUG法よりも25%小さいという結果が得られた。これは予想以上に大きな数値であり、その理由が単にデータの複雑さの計算を簡略化しただけなのか、それ以外に大きな要因が隠されているのかは、今後さらに検討していく必要がある。適用プロジェクト群の規模の大きさに差がないことから、少なくともプロジェクト規模が影響を及ぼしているとは言えない。

「その他」はIFPUG法よりもFP規模と工数の相関が低い。独自仕様という共通性の減少がばらつきを大きくしているのかもしれない。

(2) 製造業が金融・保険業よりも相関係数が高い傾向にあることについて

未調整FPはシステム特性を考慮する前のFP値である。未調整FPと工数との相関が高いということは、システム特性のバリエーションが小さいことを表していると考えられる。つまり、製造業の方が一般に金融・保険業よりもシステムのバリエーションが小さく、工数を予測しやすいものと思われる。

(3) アーキテクチャに依存しないことについて

2階層C/S、3階層C/S、ネット系では工数との相関に差が見られなかった。このことはこれらのアーキテクチャの間のシステムバリエーションの差がないと考えることができる。ただし、今回統計的に有意とはならなかったものの、2階層C/S、3階層C/S、ネット系の順で相関係数は大きくなり、この順でバリエーションがわずかながら少なくなっているのかもしれない。今後のデータ蓄積が待たれる。

(4) IFPUG法を適用したプロジェクトとNESMA概算法を適用したプロジェクトの特性の違い

今回の分析では、相関係数は適用業種(製造業と金融・保険業)によって差が見られ、適用アーキテクチャによっては差が見られないことがわかった。IFPUG法とNESMA概算法で適用業種に違いがあれば、相関係数にもその影響を考慮する必要がある。

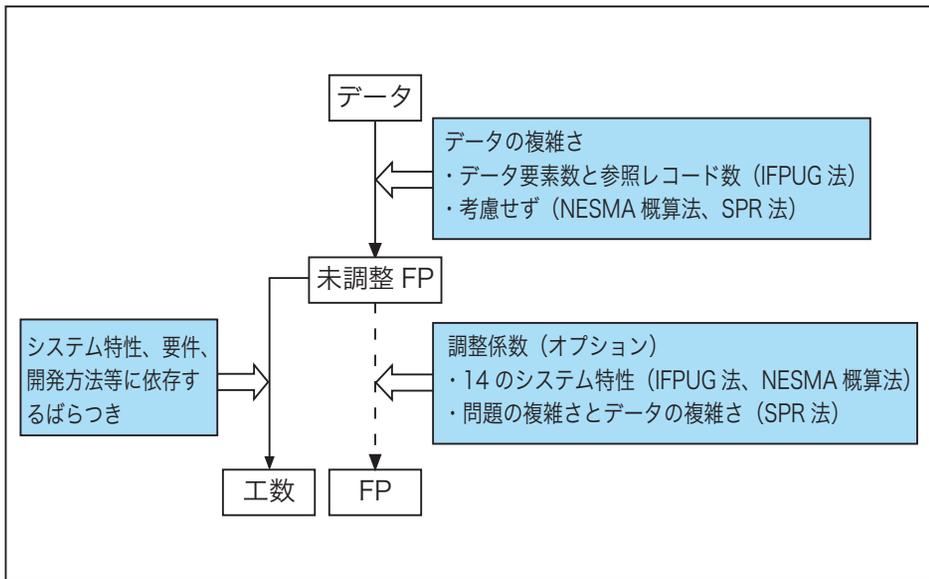


図1 FPの計算方法の概要と工数との関係

2つのFP計測手法と適用業種のクロス集計表を表6に示す。表6からIFPUG法を適用したプロジェクト群とNESMA概算法を適用したプロジェクト群の間で製造業と金融・保険業の差があるかどうかを χ^2 検定で調べたがP値は0.21となり、有意とはならなかった。すなわち、IFPUG法とNESMA概算法の差に業種の影響はないと言える。

表6 適用業種別のクロス集計表

業種 FP計測手法	製造業	金融・保険業	合計
IFPUG法	34	56	90
NESMA概算	10	28	38
合計	44	84	128

6. まとめ

(1) 層別プロジェクト件数とプロジェクトの大きさ

- 開発5工程のプロジェクトを開発種別及びFP計測種別で層別した場合、30件以上のデータが揃っているのは、新規開発と改修・保守の各4つのFP計測手法の合計8つの層から、改修・保守のNESMA概算法の層を除いた7つの層だけである。
- FP規模と工数それぞれで各層のプロジェクトの大きさを比較すると、FP規模と工数で大きさの順序は変わらない。
- 新規開発・IFPUG法と新規開発・NESMA概算法が適用されるのは最も大きなプロジェクト群で、2つのグループの間にはFP規模・工数のいずれにも有意な差は見られない。

(2) 相関分析

- 新規開発では、IFPUG法とNESMA概算法で適用するプロジェクトの大きさ及び適用業種に差がないにも拘らず、相関係数はIFPUG法の方が大きい。決定係数を比較すると、IFPUG法では0.784である(工数の78%がFP規模で決定される)のに対してNESMA概算法では0.534でその差は0.25(25%)ある。
- 同じIFPUG法を適用しても、改修・保守では新規開発に比べて決定係数が0.139(14%)小さい。

- 新規開発のSPR法は新規開発のIFPUG法よりも決定係数が0.169(17%)小さい。

(3) 業種別・アーキテクチャ別分析

- 製造業と金融・保険業の相関係数を比較すると、新規開発・IFPUG法では、製造業が金融・保険業よりも大きい傾向が見られ、「その他」では、製造業が金融・保険業よりも大きい。
- アーキテクチャ間では相関係数に有意な差は見られない。

今回の分析方法は、FP規模と工数の相関係数及び決定係数から求めたもので、同じソフトウェアを2つの方法で計測して比較したものではない。今後は相関係数以外の指標についても調査し、それらの間の関係や既存の報告などとの整合性を検討していきたい。

今回はデータ数が30以上のものに限定して相関分析を行ったが、業種別・アーキテクチャ別に層別した場合、データ数が25～29までの層が4つ、20～24の層が6つある。今回の分析ではこれらの層を分析対象外としたが、これらの層のデータ数が30に達するとさらに精度の高い分析が可能となるとと思われる。

【参考文献】

- [IPA2012] 独立行政法人情報処理推進機構 (IPA) ソフトウェアエンジニアリングセンター (SEC) 監修: ソフトウェア開発データ白書 2012-2013, 2012年.
- [JISX0142] JISX0142 ソフトウェア技術—機能規模測定—IFPUG機能規模測定手法 (IFPUG 4.1版未調整ファンクションポイント) 計測マニュアル, 2010年.
- [富野 2010] C. Jones 著、富野 壽、小坂恭一監訳: ソフトウェア開発の定量化手法 (第3版)、共立出版、2010年.
- [NESMA] Early Function Point Counting 開発初期段階でのFPカウント方法 (<http://www.nesma.nl/sectie/home/>)