

# 2P-06 オブジェクト指向メトリクスの分布モデルに基づいた 品質評価基準値の設定手法

野中 誠 福原 綾介 菅野 亨太 東 基衛  
早稲田大学大学院理工学研究科 経営システム工学専門分野

## 1 はじめに

ソフトウェアメトリクスを用いてソフトウェアの内部品質を定量的に評価し、評価基準値に基づいてプログラムを改善したりコードレビューを重点的に行うことは、ソフトウェアの品質向上に有効である。しかし、適切な評価基準値はアプリケーションの種類や個人によっても異なるため、唯一の評価基準値ですべての対象を評価するのはあまり適切ではない。また、この基準値は恣意的に決定されるのではなく、メトリクス値の分布と過去の欠陥記録との関連に基づいて決定されるべきである。

本報告では、オブジェクト指向メトリクスに関する評価基準値の設定を、アプリケーション毎あるいは個人毎のメトリクス値の分布モデルと欠陥記録に基づいて設定する手法を説明する。また、この手法を PSP (Personal Software Process) [1] の演習課題で得られたデータに適用した結果についても報告する。

## 2 アプリケーション毎のメトリクス値の分布

予備調査として、(A) 通信プロトコルを実装したデーモンプログラムと、(B) 在庫・販売管理アプリケーションを対象に、メソッド毎のステートメント数 (M-STMT)、メソッド毎のサイクロマチック数 (M-CYCL) を計測し、これらの分布を比較した。最大値とメジアン (中央値) を表 1 に、図 1 と図 2 に各メトリクス値のヒストグラムを示す。いずれの分布も大きく左に歪んだものとなった。

ここで、平均値ではなくメジアンを用いているのは、図 1 や図 2 ような分布の場合、一般にメジアンの方が代表値として優れているためである。

表 1 メトリクス値の比較 (上段:最大値 下段:メジアン)

	Total LOC	Total STMT	M-STMT	M-CYCL
(A)	294	150	24 3.0	8 1.5
(B)	2,563	1,479	52 1.0	17 1.0

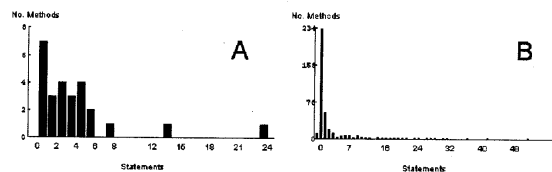


図 1 各アプリケーションの M-STMT の分布

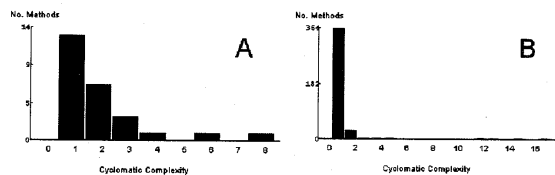


図 2 各アプリケーションの M-CYCL の分布

ここで、M-STMT と M-CYCL を計測した理由は、これらはいずれもメソッドの複雑度を表すメトリクスであり、これらの値が大きいくほど論理エラーが発生しやすくなると思われるためである。

(B) の M-STMT は (A) に比べて M-STMT が 1 のメソッドが多い。また (B) の M-CYCL 分布は (A) に比べて小さな値に極端に偏っている。これは (B) には複雑な計算を行うメソッドが少なかったためである。

このように、メトリクス値の分布はアプリケーションによって異なることが確認できた。また、欠陥記録などの品質に関するデータとの関連を調べなければ、メトリクス値について適切な評価を行うことができない。

## 3 品質評価基準値の設定手法

ここでは、メトリクス値の分布モデルと欠陥記録に基づいて、ソフトウェア内部品質の評価基準値を設定する手法について説明する。

A Method for Defining Quality Rating Criteria with Object-Oriented Metrics Distribution Model

NONAKA Makoto, FUKUHARA Ryosuke, KANNO Kyota, and AZUMA Motoei

Graduate School of Science and Eng., Waseda University

### (1) 各メトリクス値のヒストグラムを作成する

これまでに作成したプログラムの各メソッドについて M-STMT と M-CYCL を計算する。各メトリクスについてヒストグラムを作成する。

### (2) 論理エラーによる欠陥記録を関連づける

論理エラーに起因する欠陥は、PSP の欠陥記録ログのうち、欠陥を作り込んだ工程が設計あるいはコーディングで、除去工程がテストであり、且つ欠陥タイプが「機能欠陥」である欠陥データがこれに相当する。これらの欠陥データを収集する。

欠陥を含むクラスの M-STMT および M-CYCL の最大値を、各ヒストグラム上にプロットする。

### (3) 欠陥数の比率に基づいて評価基準値を設定する

各ヒストグラムにプロットされた欠陥を、メトリクス値の大きい方から順に数え上げ、全欠陥数に対する任意の比率（例えば 50%、80%）に達した時点のメトリクス値を、品質評価基準値に定める。

## 4 適用例

この手法を、PSP 演習課題 1A-5A で得られた 2 名のデータに適用した。両者の各メトリクス値の最大値とメジアン（中央値）および論理エラーに起因した欠陥数を表 2 に、図 3 と図 4 に各メトリクス値のヒストグラムを示す。

表 2 メトリクス値（上段:最大値 下段:メジアン）

	Total LOC	Total STMT	M-STMT	M-CYCL	欠陥数
(A)	352	193	22 1.0	13 1.0	7
(B)	724	411	42 4.0	16 2.0	12

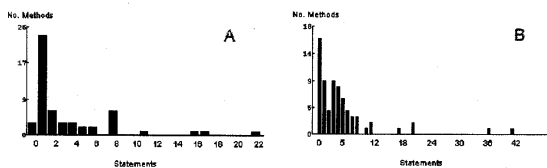


図 3 両被験者の M-STMT の分布

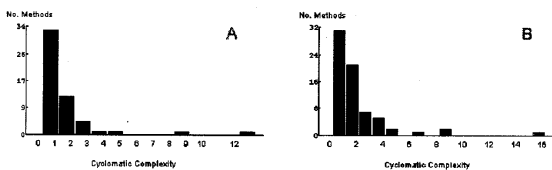


図 4 両被験者の M-CYCL の分布

### 4.1 欠陥数に基づく評価基準値の設定

これらのデータに本手法を適用し、評価基準値を計算する。表 3 に、両被験者の欠陥記録から得られた評価基準値を示す。欠陥数の比率は上段に 50%、下段に 80% の値を示した。また括弧内は、得られたデータのうち、この評価基準値以上の値を持つクラスに欠陥が含まれていた割合を表している。

ここで得られた結果の範囲では、M-CYCL の方が、論理エラーに起因する欠陥を含みやすいクラスを適切に判定できると思われる。

表 3 品質評価基準値（上段:50% 下段:80%）

	M-STMT	M-CYCL
(A)	11 (75%) 8 (60%)	5 (100%) 3 (75%)
(B)	21 (75%) 12 (57%)	9 (100%) 5 (67%)

### 4.2 欠陥修正時間との関連

修正時間と分布モデルとの関係を定性的に分析したが、これらの間には明示的な関連は見られなかった。欠陥修正時間には、分布モデル以外の要素が影響していると考えられる。

### 4.3 確率分布モデルの推定

分布モデルが確率分布に従うことが推定できれば、あるメトリクス値のメソッド発生確率の理論値が算出でき、確率分布に基づいた評価も可能になる。両被験者の分布モデルについて推定した結果、有意水準 5% で M-CYCL は指数分布として推定可能であるという結果が得られた。指数分布の確率密度関数は、

$$f(x) = \lambda e^{-\lambda x} \quad (1)$$

であり、 $\lambda$  は A が 0.530、B が 0.424 と推定された。

## 5 考察および今後の課題

今回の適用例では、被験者ごとに異なった評価基準値が得られた。このような適用実験を多くのアプリケーションや被験者に実施することで、より適切な評価基準値が得られると思われる。

また、M-CYCL の分布は指数分布に従うと推定できた。このような確立分布を評価基準値の設定などにどのように活用するかは今後の課題である。

### 参考文献

- [1] Humphrey, W. S., *A Discipline for Software Engineering*, Addison-Wesley, 1995.