

複雑さの尺度に基づく欠陥モジュールの 判別方法

3K-2

高橋 良英 若山 博文
(NTT情報通信網研究所)

1. はじめに

計画的なソフトウェアの品質管理を行うためには、開発プロセスのできるだけ早い段階でのモジュール品質の判別が必要となる。本検討では、ソースコードの複雑さから欠陥モジュールを判別する判別分析モデルにおいて⁽¹⁾、判別効率を更に向上するための、判別群境界の設定方法を提案する。群境界は、Wilksの Λ 統計量⁽³⁾を応用して設定する。これより、欠陥数(fault)の大小によりモジュールを2群に分ける場合の群境界設定の曖昧さを排除できると共に、プログラム分割等のための、複雑度の合理的基準値を見いだすことができる。本判別モデルの妥当性を実験により検証している。

2. ソフトウェア品質判別モデル

当品質判別モデルは、複雑度により、モジュール品質に傾向の違いがあることに着目し、マハラノビスの汎距離が最小となる群に、標本の所属群を決定する方法である。マハラノビスの汎距離は、各群の複雑度の重心 μ_k ($k=1, 2$)と標本の複雑度 x の距離であり、 Σ を複雑度の共分散行列とすると、以下で定義される。

$$D_{k^2}(x) = (x - \mu_k)^t \Sigma^{-1} (x - \mu_k)$$

2.1 複雑さを測る尺度 (表1)

ソフトウェアの複雑さを測る基本的尺度として、規模 (SLOC: ソースコードの行数)、制御構造、データ構造、データの流れの複雑さ、モジュール化の効用、読解性の観点から14の尺度とした。

2.2 品質を測る尺度

表1. 複雑さを測る尺度

複雑さ (略)		説明
規模	総規模 :VOL	改造+ 流用
制御の流れ	サイクロマチック数:CYC	条件文の数+1 (McCabeの尺度)
	ネスト数 :NST	called-called 関連木の深さ
処理	オペレータ出現数 :N ₁	・呼び出す関数の出現回数 ・システム共通ルーチンの使用回数 ・予約語の使用回数
	グローバル変数数 :GN ₁	global, static 各変数定義数の和
	(同上) 出現数 :GN ₂	同上の定義数, 参照回数, 更新回数の和
データ	ローカル変数数 :LN ₁	local 変数定義数
	(同上) 出現数 :LN ₂	同上の定義数, 参照回数, 更新回数の和
情報の流れ	引き数 :P	関数の引渡しパラメータ
	引き数 ² :P ²	同上の2乗
モジュール化	関数再利用回数 :cal	当関数が他の関数より呼ばれる回数
	関数の数 :fNO	1モジュール内での関数の数
	関数規模 :fsi	総規模 / 関数の数
読解性	コメント行数 :COM	コメント行数

コンパイル完了後から試験工程以降に発生したプログラムの欠陥数を、品質を測る尺度とした。欠陥数の大小で、モジュールを2群に分ける。

2.3 最適群境界モデル

欠陥数の大小によりモジュールを2群に分ける場合、どこに境界値を引けばよいか不明な場合がある。この問題は、目的変数が定量的な場合に必ず生じる。群境界設定の根拠が曖昧だと、モジュール品質の判別に誤りを生じ、試験・保守稼働に大きな影響を与える。本検討では、複雑度が欠陥に最も大きな影響を与えるという仮定のもとに、欠陥数の最適な境界値の設定方法について検討した。最適群境界モデルの決定にあたっては、群間変動の群内変動に対する割合が最も高いモデルが、予測の精度が最も高いので、Wilksの Λ 統計量を用

A Discriminant Analysis for Detection of Fault-Prone Modules Using Software Complexity Measures

Ryouei Takahashi, Hirofumi Wakayama

NTT Network Information Processing Systems Laboratories

Yokosuka-shi, Kanagawa, 238-03, Japan

いた検定により最適な群境界判定を行うこととした。アルゴリズムは以下のように単純である。先ず、群境界の初期値を0として、1ずつ追加していく。各境界モデルの判別効率Fを求める。Fは、群内変動に対する群間変動の割合で、pを説明変数数、nを標本数とすると、WilksのΛにより、 $F = ((1-\Lambda)/p) / (\Lambda/(n-1-p))$ で近似する⁽³⁾。次に、最もF値の高いモデルが最適と判断する。

3. 実験結果

3.1 実験空間

UNIX上の応用プログラム(AP)をバージョンにより2群に分ける。群AP₁は、25モジュール構成(最大2.7KS, 最小12S, 平均1KS), 群AP₂は、53モジュール(1.6KS, 16S, 360S)である。群AP₁によりモデルを構成し、群AP₂で評価した。

3.2 複雑度の主成分(表2)

14個の複雑度間は互いに相関が高い。モデルの判別精度を向上させるために、相関の高い要因をグループ化する必要がある。グループ化は主成分分析で行った(表2)。固有値の大きい2つの主成分(累積78%)を判別に効く主成分とした。第1主成分は、規模、第2主成分はモジュール化率と解釈できる。

表2. 複雑度主成分

主成分	第1	第2
解釈	規模	モジュール化
固有値	9.11	1.86
VOL	0.97	-0.22
CYC	0.96	-0.21
NST	0.62	-0.23
構造	0.96	-0.19
Gn ₁	0.43	-0.44
GN ₂	0.79	-0.38
Ln ₂	0.92	0.01
LN ₂	0.91	-0.07
P	0.90	0.39
P ²	0.73	0.64
fNO	0.95	0.22
cal	0.61	0.72
fsi	-0.08	-0.40
COM	0.93	-0.21

⋯: 主成分と相関高い

3.3 最適群境界モデル(表3)

抽出した2つの複雑度の主成分で判別分析した。

(1) 判別効率(F値)

群境界(欠陥数)と判別効率(F値)及び正答率の関係を表3に示す。この結果、欠陥が2件以下か否かで群を判別するモデルM3が、最も判別効率が高いことがわかった。群境界が0のモデル(M1)とF値にほとんど差はなく、モデルの優劣はつけられない。M3の正答率(判別結果と一致した標本の割合)が84%(=21/25)と最も高かった。同じ欠陥数でも、規模の大小により、品質の判断が

異なるので単位規模当たりの欠陥数と複雑度の関係を調べたが、判別効率は低かった(F=2.99)。

(2) 予測モデルとしての評価

AP₂に適用した結果、正答率98%(52/53)と、モデルの見積り精度は高かった。

3.4 品質管理への応用(表3)

プログラム分割基準を求めるため、分散分析法を用いて、1関数当たり(約100ステップ)のサイクロマチック数の基準値(群平均の中間値)を調べると、各モデル毎に複雑度の基準値が微妙に異なる(表3)。最適群境界モデルの基準値は11(AP₂は10)で、文献(2)の値10より高めであった。

表3. モデルの評価 ⋯: 最適群境界モデル

モデル ID	欠陥数 (群) 境界値	判別効率 F	正答率			サイクロマチック数 基準値
			1群	2群	全体	
M1	0	8.218	11/14	8/11	19/25	9.29
M2	1	4.764	13/17	6/8	19/25	10.05
M3	2	8.223	16/19	5/6	21/25	10.72
M4	8	5.520	15/20	4/5	19/25	10.92
M5	9	1.073	15/22	2/3	17/25	11.74

<注>○: 有意水準5%で群間差ありと検定(群境界=3.443)

4. おわりに

複雑さから欠陥モジュールを判別する判別分析モデルにて、群境界設定の曖昧さを除く最適モデルを、WilksのΛを応用して設定する方法を述べた。実験では、約80%の正答率となっているが、群内変動が大きい。このため、識別群数を3群(品質良、普通、不良)に分けて判別し、良・不良ソフトを分別する等の精度向上が望まれる。本最適群境界モデルを応用し、サイクロマチック数の基準値の決定等、品質管理に应用することができる。

[参考文献]

- (1) J. C. Munson & T. M. Khosgoftaar: "The detection of fault-prone program", IEEE SE., VOL. 18, NO. 5, pp. 423-433, 1992-5.
- (2) T. J. McCabe: "A complexity measure", IEEE SE., VOL. SE-2, NO. 4, pp. 308-320, 1976-12.
- (3) C. R. ラオ(奥野忠一ほか訳): "統計的推測とその応用", pp. 502-503, 東京図書, 1977.