

類似性に基づくソフトウェア信頼性予測における ファンクションポイント詳細情報の効果

吉上 康平 角田 雅照
近畿大学理工学部情報学科

1 はじめに

近年、ソフトウェアはより大規模化しており、納期の遅れ、品質の低下などを避ける重要性が増している。ソフトウェアの欠陥数をプロジェクトの早期に予測し、適切なテスト計画を立案することは、それらの問題を避けるために重要な活動のひとつである。本研究では、欠陥数の予測精度を高めるために、予測モデルの説明変数としてファンクションポイント (FP) の詳細情報を用いることを試みる。

2 ファンクションポイントによる予測

ソフトウェア欠陥数の予測値は、ソフトウェアの規模などに基づいて算出される。規模はファンクションポイント法に基づくことが多く、FPの各要素 (EI, EO, EQ, ILF, EIF) の数値を最初に求め、その数値を合算することによりFPが求められる。数学的モデルによる予測では、ファンFPの各要素を合算した後の数値を規模として用いることが多い。

ただし、工数見積もりモデルにおいて、各要素のまま説明変数に用いるほうが、合算後の数値を用いるよりも精度が高いことを示した研究[1]が存在する。本研究は、この先行研究に基づき、信頼性の予測においても同様に各要素を説明変数に用いることにより、予測精度が高まるかを確かめる。

予測方法として類似度に基づく予測方法 (CBR) を用いる。本研究では、類似度計算にピアソン相関係数とユークリッド距離を用いる。前者の場合、FPの各要素の傾向が類似しているものが、後者の場合、各要素の値に近いものが類似しているとみなされる。

3 実験

FPの各要素を説明変数として用いた場合と用いない場合、及び、類似度に相関係数を用いた場合とユークリッド距離を用いた場合の4つを組み合わせ、欠陥数を予測した。類似ケース数は8とした。また、欠陥数の平均値を単純に予測値とした場合と予測精度の差を確かめた。

分析ではISBSGデータセットを用いた。モデル構築において、開発種別、未調整FPおよびFPの各要素、開発種別、アーキテクチャ、開発言語種別を説明変数として採用した。欠損値を含まないなどの条件でプロジェクトを抽出した結果、実験に用いるケース数は134件となった。

モデルの予測精度を評価するために、リーブワウンアウト法を適用し、評価指標として絶対誤差、

Balanced Relative Error (BRE) それぞれの平均値、中央値を用いた。AEは絶対誤差、BREは相対誤差を示し、それぞれの指標の値が小さいほど予測精度が高いことを示す。

実験結果を表1に示す。要素の列はFPの各要素を用いているかを示す。欠陥数の平均値を予測値とした場合、絶対誤差は他の方法と大きな差はなかったが、その他の評価指標は大きな差があった。よって、少なくとも平均値を用いるよりは、類似性に基づく方法のほうが、予測精度が高くなるといえる。

類似度計算がユークリッド距離の場合、要素を用いたほうが、BRE中央値が大きかった。ただし、類似度計算に相関係数を用いた場合、要素を用いたほうが、精度を用いない場合よりも全ての評価尺度の値が小さく、かつ、その他の予測方法よりも評価尺度の値が小さかった。よって、類似度計算に相関係数を用い、かつFPの各要素を用いることにより、最も予測精度を高めることができるといえる。

4 おわりに

本研究では、類似性に基づく方法によりソフトウェア欠陥数を予測する際に、ファンクションポイントの各要素を用いることにより予測精度が改善することを示した。

謝辞 本研究の一部は、文部科学省科学研究補助費 (基盤C: 課題番号 25330090) による助成を受けた。

参考文献

- [1] L. Buglione, and C. Gencel, "Impact of Base Functional Component Types on Software Functional Size Based Effort Estimation," Proc. international conference on Product-Focused Software Process Improvement, 2008.

表1 各手法の予測精度

類似度計算	要素	絶対誤差	絶対誤差	BRE	BRE
		平均値	中央値	平均値	中央値
相関係数	Y	11.6	4.1	288%	146%
相関係数	N	11.9	5.0	310%	176%
ユークリッド距離	Y	13.4	4.5	349%	175%
ユークリッド距離	N	14.5	6.3	418%	158%
平均値による予測	N	14.6	10.0	479%	333%