

OSS開発における 共進化プロセスの理解のため の遅延相関分析

山谷 陽亮, 大平 雅雄, 伊原 彰紀

2013情報処理学会研究報告

背景

- OSSの進化が著しい

- ✓ コミュニティに参加する参加者が増加
- ✓ アーティファクトにおけるバグの増加
- ✓ 開発者のレベル上昇

お互いが影響しあいながらOSSは進化していくもの



共進化

遅延相関分析

相関分析

一方のレイヤーの変化が他方のレイヤーの変化に影響を与えるかを調査

遅延相関分析

どれだけの期間の遅延をもって一方のレイヤーの変化が他方のレイヤーの変化に影響を与えるか

遅延相関分析

三つのパラメータを定義

加算係数・・・説明変数の値を累積させた期間

差分係数・・・目的変数の値の変化を考慮する期間

遅延係数・・・説明変数の累積が目的変数の変化に影響を与える期間

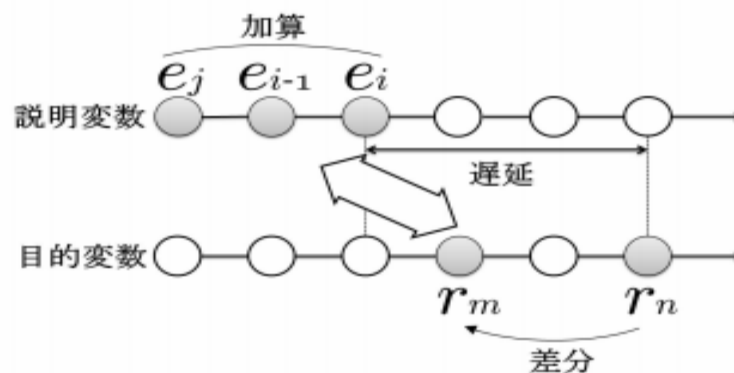


図 2 時系列データ処理の概念 ([15] を参考に改編)

アルゴリズム

STEP1 : 説明変数の累積値を求める(加算係数 $i-j$ 間)

$$e_{ij} = e_i + e_{i-1} + \cdots + e_j$$

e_i : 時刻*i*における説明変数の値

e_{ij} : 説明変数の累積値

アルゴリズム

STEP2 : 目的変数の値の変化値を求める

$$r_{nm} = r_n - r_m$$

r_n : 時刻nにおける説明変数の値

r_m : 時刻mにおける説明変数の値

アルゴリズム

STEP3 : STEP1,STEP2より相関係数を求める

→ ピアソンの積率相関係数を用いて求める

アルゴリズム

STEP4 : STEP1からSTEP3までの処理を各パラメータの
全組み合わせについて行う

STEP5 : 相関係数の絶対値が最大となる際のパラメータと
相関係数の絶対値の値を求める

対象プロジェクト

- ✓ Eclipse Platformプロジェクト
 - 2004年8月から2006年6月のデータ

メトリクス

表 1 各レイヤーにおけるメトリクス

Table 1 Metrics in each layer

進化に関するレイヤー	メトリクス名	メトリクスの説明
アーティファクト	AddLine	追加した行数
アーティファクト	AvgCyclomatic	各メソッドの複雑度の平均
アーティファクト	CountDeclClass	クラス宣言の数
アーティファクト	CountDeclFunction	関数宣言の数
アーティファクト	CountLineBlank	空白の行数
アーティファクト	CountLineCode	総コード行数
アーティファクト	CountLineCodeDecl	宣言部のコード行数
アーティファクト	CountLineCodeExe	処理を行うコード行数
アーティファクト	CountLineComment	コメントの行数
アーティファクト	CountStmtDecl	宣言部のステートメント数
アーティファクト	CountStmtExe	処理を行うステートメント数
アーティファクト	NumberOfRevisions	ファイルを改訂した回数
開発者	Committer	ソースコードを書き換えた人数
開発者	Reporter	バグを報告した人数
開発者	Fixer	バグを修正した人数
コミュニティ	NumberOfForums	Eclipse Community Forums に投稿された回数
コミュニティ	NumberOfPeople	Eclipse Community Forums に投稿した人数

結果

- ✓ メトリクスの組についてパラメータ変化させ分析
→ 48通りのパターン

結果

表 3 開発者 → アーティファクトでの結果

Table 3 Developpers → Airtifact

説明変数	目的変数	加算係数	差分係数	遅延係数	相関係数
Committer	AvgCyclomatic	3	3	3	0.533
Committer	CountLineCode	2	3	0	0.491
Committer	CountLineCodeDecl	2	3	0	0.533
Committer	CountStmtDecl	2	3	0	0.577
Reporter	AddLine	3	0	2	-0.860
Reporter	AvgCyclomatic	3	0	0	0.541
Reporter	CountDeclFunction	3	3	0	-0.583
Reporter	CountLineBlank	3	3	0	-0.421
Reporter	CountLineCode	3	3	1	-0.505
Reporter	CountLineCodeDecl	3	3	2	-0.603
Reporter	CountLineCodeExe	3	3	0	-0.547
Reporter	CountStmtDecl	3	3	2	-0.620
Reporter	CountStmtExe	3	3	0	-0.532
Reporter	NumberOfRevisions	3	3	2	-0.713
Fixer	AvgCyclomatic	3	3	2	0.490
Fixer	CountLineCodeDecl	2	3	0	0.454
Fixer	CountStmtDecl	2	3	0	0.482
Fixer	NumberOfRevisions	2	0	0	0.741

加算係数

1 ~ 3

差分係数

0 ~ 3

遅延係数

0 ~ 3

結果

表 5 コミュニティ → 開発者での結果

Table 5 Community → Developpers

説明変数	目的変数	加算係数	差分係数	遅延係数	相関係数
NumberOfForums	Committer	1	3	1	0.495
NumberOfForums	Reporter	3	3	2	0.419
NumberOfForums	Fixer	3	2	1	0.648
NumberOfPeople	Committer	1	3	1	0.421
NumberOfPeople	Fixer	1	0	2	0.443

まとめ

- ✓ OSSシステムとコミュニティは共進化の関係にあることを考慮することにより発見できた.
- ✓ Eclipse Platformだけでなく, NetbeansやFirefoxなども対象にする.
- ✓ メトリクスの追加

所感

- ✓ オフショア開発の研究よりも現実的な研究だと感じた.
- ✓ STEP4とSTEP5の動作の詳細を見てみたいと思った.