

論文内容の要旨

博士論文題目 開発履歴を用いたソフトウェア開発プロジェクトの可視化

氏 名 大蔵 君治

(論文内容の要旨)

情報化の進む現代社会において、ソフトウェアはインフラの重要な構成要素となっている。しかしながら、ソフトウェア開発プロジェクトでは今もなおコスト超過や納期遅れといった“失敗”が相次いでいる。プロジェクトの失敗が減らない理由の 1 つとして、失敗の原因を分析することが難しいという点が挙げられる。ソフトウェア開発は一般的な物づくりとは異なり、進捗（完成度）がわかりづらい、欠陥が発見しにくいといった特徴（不可視性）がある。そのため、ソフトウェア開発プロジェクトの事後分析を行うためには、開発履歴や障害管理の記録をもとに失敗の原因をプロジェクトメンバー間で精査する必要がある。しかしながら、膨大な記録を開発者の経験や勘にもとづいて個々に調べていく従来の分析手法は、時間的なコストが高く労力を要する。よって、プロジェクトメンバーの回顧を十分に促せているとはいえない。

本研究ではこのような問題に対し、ソフトウェア開発プロジェクトの「可視化」を行うことで、プロジェクトメンバー、あるいは外部の分析者に対して対象プロジェクトの事後分析を支援する手法を提案する。可視化の具体的なアプローチは、次の通りである。

- (1) ソフトウェア開発プロジェクト再現ツール「プロジェクトリプレイヤ」の開発
- (2) E メールアーカイブのクラスタリングによるソフトウェア開発コンテキストの可視化

本研究ではまず、様々なソフトウェア開発支援ツールのログデータを用いて開発プロジェクトを可視化するツール「プロジェクトリプレイヤ」を開発した。本ツールは自動的に収集された開発記録（開発ツールログ）を入力とし、コードの修正やバグ修正といった開発記録を時系列で表示することができる。必要なデータはすべて自動で収集されるため、開発者の負担を伴うことなくツールを動作させることができる。本研究では更に、「異常の原因」を発見するための手段として、開発に用いられた E メールを対象としたコンテキスト可視化手法を提案した。本手法は E メールアーカイブに含まれる種々のやりとりを文書分類技術にもとづいてクラスタリングし、他の時系列グラフと重畳表示させることで開発コンテキストの俯瞰を実現する。

本研究では、それぞれ規模・開発形態の異なる複数のプロジェクト（学生による開発・企業による開発・オープンソース開発）に対し、複数の被験者（研究者・大学院生・大学生・高専生・高校生）を用意し、(A)プロジェクトリプレイヤの有用性、(B)コンテキスト可視化手法の有用性を検証するための実験、およびケーススタディを行った。検証の結果、プロジェクトリプレイヤを用いることで、開発者に負担を掛けることなく開発プロジェクトを可視化でき、プロジェクトメンバーによる事後分析や、外部研究者による対象プロジェクトの分析を支援できることが確認できた。また、対象の規模・開発形態に関わらずプロジェクトの様々なコンテキスト（各開発者の役割、トラブルの原因等）を短時間で把握できることが確認できた。

(論文審査結果の要旨)

本論文では、ソフトウェア開発プロジェクトの事後分析支援を目的とした研究について述べている。ソフトウェアには欠陥が目に見えない、完成度が目に見えないといった「不可視性」があり、ソフトウェア開発プロジェクトを分析するにあたっては、この不可視性の問題をいかにして克服するかが課題となる。本論文ではこの課題に対し、(i) 開発記録にもとづいたソフトウェア開発プロジェクトの可視化手法、および(ii) 開発者間のやり取りに着目したソフトウェア開発コンテキストの可視化手法を提案している。本論文の成果は、以下の 2 点に要約される。

1. ソフトウェアの開発履歴を用いてプロジェクトの再現を行う可視化手法を確立し、ソフトウェア開発プロジェクト可視化ツール「プロジェクトリプレイヤ」として手法を実装した。ソフトウェア開発プロジェクトの事後分析は、一般に人的にも時間的にも多大なコストが掛かる。提案する可視化手法は、ソフトウェア開発において広く用いられている開発支援ツールが自動で記録するログデータを用いて開発プロジェクトの可視化を行っている。そのため、エンジニアに対し、事後分析のための追加作業を課す必要がなく、データ収集における人的・時間的コストを大幅に軽減できる。また、実装した可視化ツールは、極めて簡単な操作でソースコードの修正や障害報告といった多種の開発情報を時系列ベースで提示することができ、分析における時間的コストも低いと言える。
2. ソフトウェア開発において用いられる開発者間の E メールコミュニケーションから、開発現場において発生した事象や種々の打ち合わせといった開発背景(コンテキスト)を自動抽出し、時系列上にプロットする可視化手法を提案した。成果物の規模、欠陥の数等、開発記録の多くは数値で表される定量的なデータである。しかし、正確なプロジェクト分析を行うためには、数値には表れない現場の情報も必要不可欠である。提案手法では、文書分類技術を応用し、開発に用いられた E メールコミュニケーションのアーカイブから、開発者間のやり取りをトピック別に可視化している。トピックを時系列で見ることで、他の時系列チャートとの比較分析が行いやすくなり、統計データの裏付けを得る糸口となる。また、提案手法は開発者間の膨大なやり取りを要約することで、事後分析における開発者へのヒアリング作業を軽減・支援しており、現場に即した分析手法として極めて有用性が高いと言える。

以上のように、本論文は、エンジニアの不足や開発の短納期化にともなう事後評価の欠如という、近年のソフトウェア開発業界が抱える問題を見据えた上で、実学的なアプローチで解決を図っており、その有用性は高い。また、本論文では提案手法の実開発プロジェクトへの適用も示されており、導入容易性の高さから、産学連携の促進という点においてもその貢献が期待できるものである。よって、本論文は博士(工学)の学位論文としてふさわしいものと認める。