

## プロジェクト管理ツールIPA EPM-Xの機能拡張による リポジトリマイニング研究基盤の構築

大平 雅雄<sup>†1</sup> 柏 祐太郎<sup>†1</sup> 松本 明<sup>†1</sup>  
山谷 陽亮<sup>†1</sup> 吉行 勇人<sup>†1</sup>

我々の研究グループでは、プロジェクト管理ツールIPA EPM-Xを機能拡張し、リポジトリマイニングを支援するためのプラグイン環境を構築している。本ワークショップでは、リポジトリマイニング研究者のための研究基盤として構築中の支援環境をどのように利用できるかについて議論する。

### A Platform for Studying Mining Software Repositories

MASAO OHIRA,<sup>†1</sup> YUTARO KASHIWA,<sup>†1</sup> AKIRA MATSUMOTO,<sup>†1</sup>  
YOSUKE YAMATANI<sup>†1</sup> and HAYATO YOSHIYUKI<sup>†1</sup>

#### 1. はじめに

社会基盤として高度なソフトウェアシステムが広く普及した昨今、交通システムや金融システムの障害による業務・サービスの停止や中断は、我々の社会経済活動に多大なる損失を与えるようになってきている。社会活動におけるソフトウェアシステムの重要性が今後も増々高まると予想される一方で、ソフトウェア開発環境のグローバル化による競争の激化から、開発の短納期化・低コスト化が同時に進むという状況にある。このような過酷な環境下において、ソフトウェア品質の確保や納期を遵守するためには、進行中のプロジェクトをリアルタイムにモニタリング（検出不具合数や工数の進捗具合などの計測と可視化）し、プロジェクト内で発生する問題を早期に検出し、対処する必要がある。

このような社会的要請に対する取り組みとして、独立行政法人情報処理推進機構（IPA）が普及に取り組んでいるIPA EPM-X<sup>\*1</sup>がある。IPA EPM-Xは、ソフトウェア開発状況を定量的に把握しプロジェクト管理を支援するツールである。EPM-Xは、文部科学省のリーディングプロジェクト「e-Society 基盤ソフトウェアの総合開発」の一環のEASE（ソフトウェア工学へのエンピリカルアプローチ）プロジェクトにおい

て開発された、EPM (Empirical Project Monitor)<sup>1)</sup>をモデルとしている。

ただし、現状のEPM-Xは、定量的プロジェクト管理の産業界への普及に重きを置いているため、基本的な定量データ（ソース規模、工数、進捗、品質等）の自動収集と、グラフ化によるプロジェクト管理支援機能が提供されているのみである。したがって、基本的にはプロジェクト管理者がグラフ等から問題の発生を目視で発見し対策を講じるというリアクティブなプロジェクト管理にならざるを得ないという問題がある。前述したように、昨今の厳しい開発環境下においてソフトウェア開発の現場で求められているのは、プロジェクト内の異変や問題発生の予兆をリアルタイムに検出し、問題の発生を未然に防止するプロアクティブなプロジェクト管理と言える。

我々の研究グループでは、ソフトウェア工学分野の先導的研究支援事業（独立行政法人情報処理推進機構）「IPA EPM-Xの機能拡張によるプロアクティブ型プロジェクトモニタリング環境の構築—次世代の定量的プロジェクト管理ツールとリポジトリマイニング研究基盤—」の一環として、IPA EPM-Xにリポジトリマイニング機能を追加するための研究開発に取り組んでいる。次章では、現在構築中の支援環境について述べる。

<sup>†1</sup> 和歌山大学システム工学部  
Wakayama University

<sup>\*1</sup> 定量的プロジェクト管理ツール EPM-X,  
<http://www.ipa.go.jp/sec/softwareengineering/tool/ipf>

## 2. リポジトリマイニングプラグインの開発

### 2.1 目的

IPA EPM-X のプロジェクト管理支援は主に、Trac または Readmine の課題管理システム、Git または Subversion の構成管理システム（のリポジトリ）に蓄積されたデータを活用して提供される。これらリポジトリに蓄積されるデータは、近年リポジトリマイニング分野においてもソフトウェア開発の様々な支援を目的として活用されているものである。そこで我々の研究グループでは、近年ソフトウェア工学分野で進展が目覚ましい (1) リポジトリマイニング技術と、リポジトリマイニング技術を発展させた (2) プロアクティブマイニング技術を定量的プロジェクト管理ツール EPM-X に取り入れることを目的として、リポジトリマイニングプラグインの開発に取り組んでいる。実務者が手軽に最新の技術を利用できるようにすることが狙いである。

### 2.2 リポジトリマイニングプラグイン

リポジトリマイニングとは、構成管理ツール（Subversion や Git など）、プロジェクト管理ツール（Trac や Redmine など）および不具合管理ツール（Bugzilla など）に蓄積された膨大な開発履歴データベース（リポジトリ）に対するデータマイニング技術の総称であり、ソフトウェア開発に有益な知見を過去の定量データに基づいてフィードバックすることを目的としている。盛んに研究が行われている一方で、実務者が気軽にリポジトリマイニング技術を試したりプロジェクトに適用したりするためのツールは未だ不十分な状況にある。そこで我々は、トップ会議やトップジャーナル（ICSE, FSE, TSE など）で発表されたリポジトリマイニング技術を厳選し、プラグインとして実装した。現在、以下の 5 種類のプラグインが利用可能である。

- バグトリアージ
- Tossing Graph
- 不具合修正時間予測
- バグモジュール予測
- 重複コード検出

### 2.3 プロアクティブマイニングプラグイン

リポジトリマイニング技術は、蓄積された膨大なプロジェクトデータに基づいて次期プロジェクトの工数を予測したり、ある程度開発が進んだ後にバグが多く含まれているモジュールを予測する、などの用途に向いているが、進行するプロジェクトをリアルタイムに支援するためのものではない。一方、EPM-X を代表とする従来の定量的プロジェクト管理技術は、プロジェ

クトの進捗をリアルタイムに把握するなどの用途に向くが、問題が発生したことをできる限り早く検知するためのものであり、プロジェクト管理者は問題を認識してからリアクティブに対応せざるを得ない。

プロアクティブマイニングでは、従来のリポジトリマイニング手法を発展させ、より少ないデータから近い将来プロジェクトに生じる何らかの異常や予兆を検知するための技術である。ソフトウェア工学分野以外では、プラント故障の予兆検出や不正アクセスの早期検出などに利用されている。現在、以下の 3 種類のプラグインが利用可能である。

- 遅延相関分析
- 変化点検出
- 変更波及分析

## 3. リポジトリマイニング研究基盤としての活用

リポジトリマイニング技術は現在も盛んに研究が行われており、今後の発展も期待できる。その一方、各研究者が独自の改良を加え、別個のデータセットで手法の有用性を独立して検証しているため、どのプロジェクトにも適用可能な一般性の高い手法が存在するのかわからず把握できない状況にある。特に先行研究との比較実験を厳密に行いたい場合には、自身が研究開発した手法とは別に、先行研究の手法を複数実装する必要があり、また、先行研究で用いられているデータセットを用意する必要があり、現在、研究者には大きな負担がかかっている。

我々のリポジトリマイニング環境を用いれば、各研究者が研究開発した手法・技術を共通のデータセットに対して適用することができるため、手法の比較や一般性の検証が容易に行えるものと思われる。今後は、本支援環境をリポジトリマイニング研究基盤としてどのように発展させ、普及につなげるかが課題になる。

**謝辞** 本研究の一部は、文部科学省科学研究補助金（基盤 (C):24500041）による助成を受けた。また、独立行政法人情報処理推進機構が実施した「2013 年度ソフトウェア工学分野の先導的研究支援事業」の支援を受けた。

## 参考文献

- 1) 大平雅雄, 横森勲士, 阪井 誠, 岩村 聡, 小野英治, 新海 平, 横川智教: ソフトウェア開発プロジェクトのリアルタイム管理を目的とした支援システム, 電子情報通信学会論文誌 D-I, Vol.J88-D-I, No.2, pp.228-239 (2005).