
E-mail データマイニングに基づく適任開発者の推薦手法の検討

A method for recommending appropriate developers based on e-mail data mining

福井 克法* 大平 雅雄†

あらまし 本研究は、e-mail データを用いることで要件定義、設計、コーディングなど、ソフトウェア開発の各工程に関与したすべての開発者を対象とした適任開発者の推薦手法を提案する。

1 はじめに

近年、ソフトウェアのニーズが高まると同時にソフトウェア開発が大規模化している。大規模なソフトウェアを短納期で開発するためには、プロジェクトに適任の開発者を割り当てることが重要である。しかし、大規模なソフトウェア開発現場では、各開発者が今までに関わったプロジェクトを把握することは難しい。そのため、プロジェクトに適任の開発者を割り当てられず、開発を円滑に行うことができない可能性がある。また、企業が開発案件を提案する際に開発者の知識や経験を生かせず、案件を逃してしまうといった問題がある。

既存研究 [?] [?] では、ソースコードの変更履歴に着目しモジュール単位での適任の開発者を推薦する手法が提案されている。しかし、実際のソフトウェア開発では要件定義や設計の開発工程が存在する。要件定義や設計を行う開発者は、プロジェクトの全体像を把握しておく必要があるため、開発経験が豊富な開発者であることが望ましい。既存手法ではソースコードの変更履歴にのみ着目しているため、要件定義や設計などに関わっているソースコードを変更しない開発者を考慮できていない。本研究では、要件定義や設計などソースコードを変更しない開発者も含めたプロジェクト単位での適任開発者の推薦手法を提案する。

2 提案手法

ドメイン用語から特定のプロジェクトの適任開発者を推薦することを目的とする。ドメイン用語とは、特定のプロジェクト内で使われる特徴的な用語のことを指す。

2.1 データソースの選択

ソフトウェア開発において頻繁に利用されるコミュニケーションツールの一つとしてメーリングリストが存在する。メーリングリストは全工程の開発に関わっている開発者が使用するため、既存手法では考慮されていなかったソースコードを変更しない開発者も対象とすることができると考える。本研究では特に、開発者同士の e-mail でのやりとりに着目する。開発者同士の e-mail には、その開発者が今までに関わった開発に関するドメイン用語が含まれていると考えられる。開発者が e-mail で使用するドメイン用語を抽出することで、開発者が今までにどのようなプロジェクトに関わったかを把握することができ、プロジェクトに対する適任の開発者を推薦することができる。と考える。

*Katsunori Fukui, 和歌山大学

†Masao Ohira, 和歌山大学

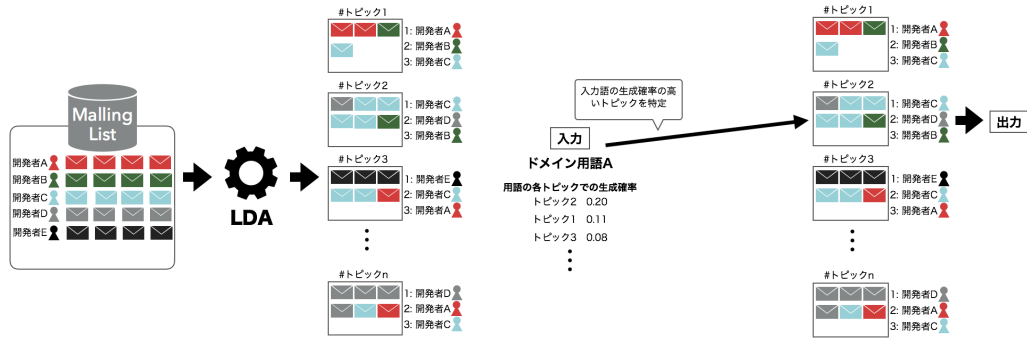


図1 ドメイン用語の抽出処理

図2 入力用語から推薦開発者を出力する処理

2.2 E-mail 分類とドメイン用語の生成確率の算出

E-mail から特定のプロジェクトのドメイン用語と関連する用語を抽出するために、LDA (Latent Dirichlet Allocation) [7] を用いる。LDA を適用した概要を図1に示す。LDA では各単語には潜在的なトピックが存在すると想定しており、各 e-mail を構成単語に応じて、指定した数のトピックに分類することができる。同じトピックに分類される e-mail は関係性が深い内容の e-mail である可能性が高い。また、LDA では各トピックの構成用語の生成確率を抽出することができる。構成用語の中にはドメイン用語や関連語も含まれていると考えられる。

2.3 開発者推薦手順

入力のドメイン用語から開発者を推薦する概略図を図2に示す。まず、2.2節で作成したトピックの中で、入力のドメイン用語の生成確率が最も高いトピックを特定する。特定したトピックには、入力のドメイン用語に関する内容の e-mail が分類されていると考えられる。次に、トピックに分類されている e-mail の件数を開発者ごとにスコアリングする。スコアが上位の開発者を入力ドメイン用語に関する開発に適任の開発者として推薦する。

3 おわりに

本研究では、要件定義や設計などソースコードを変更しない開発者も含めたプロジェクト単位での適任開発者の推薦手法を提案した。今後は複数のプロジェクトのメーリングリストを組み合わせたデータに対して提案手法を適用し、本手法の有用性を示す。また、ドメイン用語の抽出方法を改善することで精度の向上を目指す。

謝辞 本研究の一部は、JSPS 科研費(基盤 (A): JP17H00731, 基盤 (C): JP15K00101) による助成を受けた。