

## アドバンス・トップエスイー プロフェッショナルスタディ



# プログラム修正範囲特定のための設計書と ソースコード間のマッピングツールの提案

株式会社リンクレア

伊田侑起

y。ida@lincrea。co。jp

## 開発における問題点

システムの保守において、プログラム改修前に 修正範囲を調査することが多々ある。修正範囲 を調査する際にプロジェクト経験が浅い人は、 システム構成や業務・プロジェクトに関する知識 が乏しく、時間を費やしてしまう傾向がある。影 響調査方法は人によって統一されていない場 合もあり、自身の経験や勘によって調査をして



手法・ツールの適用による解決

設計書の文章とソースコードドキュメントに関連性が あると仮定し、設計書の文章とソースコードドキュメ ントを抽出し、LSIアルゴリズムを用いて設計書とソー スコード間のマッピングツール(関連表)を作成。 関連表をもとにソースコードの修正範囲を探す。 マッピングツールを使用することで画一的な方法で、 修正範囲調査の作業時間短縮を図る。

# システム概要



関連表 时, 河台, 旅, 662, 7 村, 河台, 旅, 662, 7 村, 河岳, 大川, 河岳, 高級仕入計上コントローラ ・ @author x.lshimaru (\*ログイン/、\*ユーザ、\*連 7分、74世、\*\*年、78年、\* \*コピー、7編集\*)

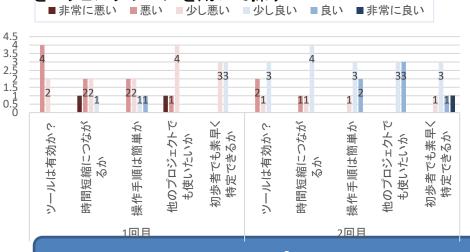
- 1。設計書の文章とソースコードドキュメントを抽出
- 2。抽出した文章に対し前処理を行い、その後分かち書きで 名詞のみにする
- 3。名詞のみとなった単語群からコーパスを作成
- 4。コーパスをtf-idf変換し、LSIアルゴリズムを適用しモデルを作成

例)「計上年月日、初回計上年月日と処理年月日が同じものを取得する」 関連表は変更した設計書の文章に関わるソースコードを関連度が高い順 に表示。表示内容はソースコードとメソッド、Javadoc。

※LSIは高次元の文書群から、与えられたクエリの意味的に 関連する文書の類似度を求めるアルゴリズム。

# 実験•評価

6人に対して実験。2回実施。2回目はモックで実験 対象者は設計書を修正した箇所と対応するソースコード をマッピングツールを用いて探す



1回目の精度が悪い状態での実験→評価が低い 2回目の精度が高いモックでの実験→評価が高い

精度が高ければ、実際のプロジェクトでも使用できる

#### ●課題

→現状の精度では実際に活用はまだ厳しい

- •対策
- (1)ソースコードのメソッド名や変数名も学習に加える →Stop word や辞書を作成しただけでは劇的な精 度向上が出来なかったので厳しいと思われる
  - ②アルゴリズムを変えてみる
    - →根本的にアルゴリズムを変更 → BERT

トップエスイー サイエンスによる知的ものづくり教育プログラム