

# 先端科学技術研究科 修士論文要旨

|   |                          |     |              |
|---|--------------------------|-----|--------------|
| 所属研究室<br>(主指導教員)  | ソフトウェア設計学<br>(飯田 元 (教授)) |     |              |
| 学籍番号  | 2311280                  | 提出日 | 令和 7年 1月 20日 |
| 学生氏名  | 森川 靖仁                    |     |              |
| 論文題目  | トークン粒度でのレビュー指摘箇所推薦モデルの提案 |     |              |
| 要旨  |                          |     |              |
| <p>コードレビューは品質保証において重要な役割を果たし、可読性、保守性の向上などソフトウェア開発において多くの恩恵をもたらすことが明らかになっている。しかし、コードレビューは多大な時間を要することが明らかとなっており、レビュアーは大量に届くレビュー要求のひとつひとつを丁寧に確認する必要がある上、レビュー対象となる変更は数十行にも及ぶ。このレビュアーに対する負担を軽減するために、多くの研究がレビュアーがレビューする必要のあるコードを推薦するアプローチを提案している。しかし、現在最も細粒度のアプローチである行レベルでの推薦を利用しても、推薦の粒度が依然として粗いため、開発者がどの箇所をレビューする必要があるかを把握することが困難な場合がある。例えば、QtBaseプロジェクトのソースコードの1行あたりに含まれるトークン数の中央値は7であり、12を超えるトークンで構成されている行は全体の10%に及ぶ。そこで本研究では、従来の研究よりも細粒度であるトークンレベルのアプローチを提案し、開発者がレビューする場所をより細かく推薦する。具体的には、CodeBERTをFine-tuningし、コメントや修正の指摘が行われる可能性のあるトークンをレビュアーに返すように設計する。OpenStackおよびQtBaseデータセットを使用した実証評価では、行レベルの評価において提案されたアプローチが行の修正を予測する最新モデルよりも良いパフォーマンスを発揮することが示された。また、トークンレベルでの評価において、修正される行を正しく予測できた場合、実際に変更されたトークンの最大86%を変更されるトークンとして本モデルが推薦できることが明らかとなった。</p> |                          |     |              |