Abstract

Software systems are developed and maintained in fast-paced, collaborative environments where effective bug management is vital for ensuring reliability and traceability. Although issue trackers and version control systems record valuable information about reported problems and corresponding code changes, jointly analyzing these artifacts remains difficult due to unstructured issue discussions, missing links, inconsistent commit documentation, and unreliable evaluation methods.

To address these challenges, this thesis presents an integrated framework for smart software bug management grounded in issue–commit analysis. It explores how textual and non-textual information from issue trackers and version control systems can enhance traceability, comprehension, and documentation across the software lifecycle. The work begins with automated methods for summarizing long issue discussions, followed by the development of robust models for issue–commit link recovery that remain reliable under adversarial and noisy conditions. Building on these foundations, a novel one-to-many issue–commit dataset and a learning-to-rank framework are introduced to capture realistic multi-commit relationships. The research further extends into commit documentation by proposing multi-level message generation for large diffs and investigates large language models as scalable evaluators of commit message quality. Collectively, these contributions form a unified and data-driven pipeline that improves understanding, linking, and evaluation in software maintenance, laying the groundwork for more intelligent and automated bug management in both open-source and industrial contexts.
Keywords: Smart bug management, Issue–commit analysis, Adversarial robustness, Issue summarization, Commit message generation, LLM-based evaluation, Software repository mining.