

テスト初心者に向けたテスト実施スキル向上手法の提案

－ “SOHT トレーニング法” の提案 －

主査 : 秋山 浩一（富士ゼロックス株式会社）
副主査 : 喜多 義弘（東京工科大学）
リーダー : 武田 匡広
研究員 : 雨宮 真実（株式会社メタテクノ）
小此木 済（株式会社インテリジェンスビジネスソリューションズ）
清水 専一（株式会社コスモコンピュータセンター）
藤野 真澄（富士通株式会社）

概要

短納期、低コストという条件下でのテスト実施では、効率よく欠陥を検出できる上級テストエンジニアが求められている。しかし、現実には経験の浅い初級テストエンジニアによりテストが実施され、必ずしも欠陥検出の効率が良いとは言えない。そこで本論文では、テスト効率の改善策として初級テストエンジニア向けのテスト実施トレーニング手法を提案する。上級テストエンジニアが欠陥検出に至った過程を欠陥情報から抽出、可視化してテスト実施に活用する一連の手順をトレーニングとする。上級テストエンジニアは豊富な経験と知識から欠陥が作り込まれやすい箇所を察知して重点的にテストしていると考え、実際の欠陥情報からテストケース以外のテストで検出された欠陥情報を用いる。トレーニング手法の内容と試行した結果について報告する。

1. はじめに

テスト工程においては効率の良いテスト、すなわち短期間、低コストでより重篤な欠陥を検出することが求められている^[1]。しかし経験の浅いテストエンジニア（以下、初級^[2]TEとする）が実施するテストは必ずしも効率が良いとは言えない。テスト効率を改善する方法としては、欠陥情報を活用してテスト設計を行い、テストケースを充実させることが考えられる。しかし、テストケースを充実させることはテストケースの肥大につながり、維持や管理が難しいという問題がある。そこで我々は、経験が豊富なテストエンジニア（以下、上級^[2]TEとする）と初級TEの欠陥検出スキルの差は欠陥を検出する過程における思考の違いにあると考え、その差を補うことに着目した。

本論文では、上級TEは豊富な経験と知識から欠陥が作り込まれやすい箇所を察知して重点的にテストしていると考え、欠陥の検出過程を「状況」と「操作」、そこから推測される「設計上の弱点」を整理した“SOHT”（ソット, Situation/Operation/Hypothesis Table: 状況/操作/仮説表）を定義し、このSOHTを用いたテスト実施トレーニング手法「SOHTトレーニング法」を提案する。

2章で研究の背景、3章で研究のねらいを述べ、4章で提案手法を説明する。5章で手法の評価結果、6章で課題を記し、最後にまとめを述べる。

2. 研究の背景（現場の課題と先行研究）

近年、ソフトウェア開発においては効率の良いテスト、すなわち短期間、低コストでより重篤な欠陥を検出するテストが必要とされ、上級TEの需要が高まっている。製品開発に占めるソフトウェアの比率は増加しており、ソフトウェアの規模も複雑さも増大している。その一方で、競争力のある製品を市場に出すためにソフトウェアにも開発期間の短縮、コストの低減が求められている^[1]。テスト工程では設計製造工程の延伸による影響に対応する必要がある。

ソフトウェアテストの現場では上級TEが不足しており、テストエンジニアの育成が課題になっている。レビューにおいてはSQiP研究会第3分科会「ソフトウェアレビュー」において長期に渡り、レビュ

第5分科会（教育チーム）

アの育成手法について研究されている。レビューオリエンテーションキット^[3]やEIDeR-Training法^[4]、DPDT法^[5]といった手法、教材などが存在する。テストにおける育成とはテスト技法の習得やJSTQB等の資格取得を指すことが多く、具体的な手法は確立されていない。

過去の研究において、欠陥情報を利用してテストケースを充実させる方法で上級TEのノウハウを初級TEに伝えることができることは確認されている。SQiP研究会第5分科会「ソフトウェアテスト」では欠陥情報からテストケースを抽出する手法について研究されてきた^{[6][7]}。SQiPシンポジウムでも欠陥情報をデータベースに整備することでテストケースを強化する研究の報告がある^[8]。いずれの場合も初級TEのテスト観点を強化することができたとの報告がされている。

しかし、GUIを持つアプリケーションなど操作の自由度が高い場合は、すべての操作を想定してテストケースを作成するのは現実的ではない。テストケースを充実させることはテストケースの肥大につながり、維持や管理が難しくなるという問題がある。また、要求仕様書や設計書に漏れや誤りが存在する場合はテストケースを充実させる方法では必ずしも対処できないという懸念がある。

効率的なテストの実施には、テストケースを充実させるだけでなく、テストケースにはないパターンを推測するスキルが必要である。上級TEは勘と経験により、テストケースにないパターンからも欠陥を検出する^[1]。SQiP研究会第3分科会「ソフトウェアレビュー」にて研究されているHDR法^[9]、DCC法^[10]や第7分科会「欠陥エンジニアリング」にて研究されている欠陥予測アルゴリズム^[11]のように、重篤な欠陥の兆候を察知するスキルを向上させることが効果的であると考えられる。

そこで我々は、従来研究されてきた欠陥情報からテストケースを作成する手法に欠陥を推測する観点を追加し、初級TEのテスト実施スキルを向上させるトレーニング手法について研究した。

3. 研究のねらい

本研究の目的は「初級TEのテスト実施スキル向上トレーニング手法」の確立である。我々は上級TEと初級TEの欠陥検出スキルの差は欠陥を検出する過程における思考の違いにあると考え、その差に着目した。上級TEはテストケースに無いパターンからも欠陥を検出することから、自身の持つ知識と経験から欠陥が作り込まれやすい箇所を推測して重点的にテストしていると考えられる。テストケースの実行で検出された欠陥は既にテストケースとして導出されており、形式知として表現されていることから、上級TEがテストケース以外のパターンで欠陥を検出した過程に暗黙知が存在すると考えた。

我々は初級TEに、上級TEが欠陥を検出する過程の思考を疑似体験させるトレーニングが有効ではないかと考えた。そこで上級TEの持つ暗黙知を表現する形式として、上級TEがテストケース以外のパターンで欠陥検出に至った過程を抽出、可視化する表を定義した。初級TEは定義に従って欠陥情報から表を作成し、表をもとに欠陥が作り込まれやすい箇所を推測してテストを実施することで上級TEの思考を辿れると考えた。

4. 提案する手法

本研究では、上級TEの有する欠陥検出に関する知識が初級TE向けの欠陥検出トレーニングに有益であるという仮説に立ち、欠陥が検出される過程を表現したSOHTを作成して欠陥検出に利用する「SOHTトレーニング法」を提案する。実際のテスト環境下でOJTの一環として実施する。

4.1. SOHTについて

SOHTはテストケース以外のパターンで上級TEが欠陥を検出する過程をまとめた表である。欠陥を検出した「状況」、そのときに行った「操作」、それによって推測できる設計上の弱点を「仮説」とする。表1にSOHTのサンプル、表2に内容を示す。SOHTはテストをする観点の一覧表であり、テストの結果より設計上の弱点を推測している点を特徴とする。推測した弱点と同一または類似の弱点を関連付けることにより、テスト対象の幅を広げるように工夫している。(図1)

第5分科会（教育チーム）

4.2. SOHTの作成手順

SOHTは実際の欠陥情報のうちテストケース以外のパターンで検出された欠陥情報から作成する。SOHTは大きくなりすぎると利用が難しくなるので、A4用紙2,3枚に収まるようにする。

表1 SOHT の例

| 状況 | 操作 | 仮説 | | |
|----------------------|---|----|-----------------------|----|
| | | ID | 推測される弱点 | 関連 |
| モニタ解像度の変更が認められている。 | SVGで画面サイズを設定したのち、XGAなどでより小さい画面で起動する。 | 1 | モニタ解像度の違いを考慮した設計が不十分。 | 2 |
| | 縦画面・横画面（ピボットモニター）を切り替えてアプリケーションを運用する。 | | | |
| | ウィンドウサイズを記憶している場合、モニタ解像度変更時の動作が考慮されているかを確認する。 | | | |
| DPIの変更が認められている。 | 72DPIに設定しアプリケーションを運用する。 | 2 | モニタ解像度の違いを考慮した設計が不十分。 | 1 |
| | 92DPIに設定しアプリケーションを運用する。 | | | |
| キーボード（KB）での操作が可能である。 | マウスを使わずKBだけで操作する。 | 3 | コントロール上のフォーカス管理が不十分。 | - |
| | チェックをマウス、OKをKBで操作する。 | | | |
| | チェックをKB、OKをマウスで操作する。 | | | |

表2 SOHT の内容

| 項目 | 内容 | |
|----|--------------------------------------|---|
| 状況 | そのテストを実施したときの状況。どういう前提で「操作」をしたのかを表す。 | |
| 操作 | どのような操作をしたのか。 | |
| 仮説 | ID | この仮説のID。 |
| | 推測される弱点 | 状況＋操作より、設計上のどのような弱点が考えられるか。内部設計を知らないテスターでは判断がつかないことがあるので、その場合は開発者に確認する。 |
| | 関連 | 推測される弱点が同一または類似の仮説ID。 |

| 状況 | 操作 | 仮説 | | |
|-----|-----|----|---------|----|
| | | ID | 推測される弱点 | 関連 |
| XXX | xxx | 1 | ... | 3 |
| YYY | yyy | 2 | ... | |
| ZZZ | zzz | 3 | ... | 1 |

状況XXXでxxxの操作をして欠陥を検出した場合、同様の弱点を有するZZZでzzzをしても、欠陥を検出できる可能性が高いと考える。

図1 SOHT を利用した類似弱点の欠陥検出

第5分科会（教育チーム）

以下にSOHTの作成手順を示す。

- (1) 要約
実際の欠陥情報から欠陥の検出手順(再現手順)を一言に要約する。要約は「状況XXXのときに操作xxxをしたら・・・となった」の形でまとめる。要約の例を表3に示す。
- (2) グループ化
状況XXXが共通するものを同一状況と考えてグループ化し、「操作」を一箇所に集約する。
- (3) 弱点の関連付け
「操作」から「推測される弱点」を洗い出し、類似の弱点を関連付ける。「・・・」の部分は多くの場合、「期待外の動作となった」など同様の表現となる。

表3 要約の例

| 再現手順 | 要約 |
|--|------------------------------------|
| 【前提条件】 解像度がXGAの状況で管理画面を開く 【手順】 1.管理画面を開く。 2.画面内のリストAのヘッダをクリックする ⇒管理画面のスクロールがユーザーの意図しない動きをする 【期待値】 管理画面のスクロールバーが移動しないこと | 画像解像度を変更した状況で、GUIを操作したら期待外の動作となった。 |

4.3. SOHTを用いたトレーニング(SOHTトレーニング法)

図2に「SOHTトレーニング法」の手順について示す。次の4ステップで実施する。

- (1) SOHTの準備(作成)
上級TEが検出した欠陥のうちテストケース外で検出された欠陥について、初級TEは初のSOHTを作成する。SOHTは上級TEの指導の下で初級TEが個人で作成する。上級TEの指導は主に初SOHT作成時とし、以降は個人の状況に応じて指導する。
- (2) テスト実施
初級TEは自身が作成したSOHTを見ながらテストを実施する。テスト対象がSOHTの「状況」と同じ局面にある場合、「操作」に示された手順を試す。ここで欠陥を検出した場合、同じ「状況」下の他の「操作」も確認する。そこから推測される「仮説」と関連する「仮説」にも着目し、同様に「操作」を試す。
- (3) SOHTの評価
初級TEは実際に欠陥を検出することができた「仮説」を記録する。欠陥を検出した「仮説」は有益な「仮説」であると考える。
- (4) SOHTの更新
ステップ(3)の評価を踏まえて、初級TEは自身のSOHTを更新して次のテスト実施に備える。テスト実施中に自分で気がついた「状況」、「操作」、「仮説」があれば追加する。欠陥を検出することができていない「仮説」は必要かどうか検討し、不要であれば削除する。

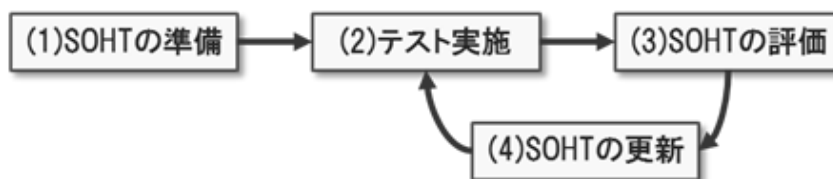


図2 SOHTトレーニングの手順

第5分科会（教育チーム）

5. 手法の評価

提案手法の評価方法およびその結果について述べる。

5.1. 評価方法

テスト実施初心者4名を被験者として選定し、2名ずつの2チーム(チームA/チームB)に分けた。テスト対象機能として開発規模、複雑度が適度に異なる4機能(機能1～機能4)を用意した。(表4)

表4 テスト対象機能一覧

| 機能名 | 開発規模 | 欠陥検出密度 | GUI |
|-----|----------|------------|-----|
| 機能1 | 約10 KLOC | 3.5 件/KLOC | なし |
| 機能2 | 約20 KLOC | 10 件/KLOC | あり |
| 機能3 | 約25 KLOC | 10 件/KLOC | あり |
| 機能4 | 約40 KLOC | 15 件/KLOC | あり |

各テスト対象機能については要求仕様書およびテスト仕様書を与え、被験者はこれらのドキュメントを自由に参照できるものとした。テスト実施順序が欠陥検出率に影響する点を考慮して、チームAとチームBでテストの実施順序を入れ替えた。

- ・チームA
機能1, 機能2に対して従来どおりのテストケースベーステストを実施した後に、機能3, 機能4について提案の手法を適用したテストを実施
- ・チームB
機能3, 機能4に対して従来どおりのテストケースベーステストを実施した後に、機能1, 機能2について提案の手法を適用したテストを実施

テスト実施の制限時間は設けず、終了判断は各チームに任せた。被験者がSOHTにある「仮説」を十分に試すことによりテストケースにないパターンの推測力が養われることを確認するためである。検出した欠陥はその重篤度合いに応じて重/中/軽の3段階に分類した。(表5)

表5 欠陥重篤度の判断基準

| 重篤度 | 定義 | 具体例 |
|-----|--|--|
| 重 | ・ユーザーが著しい被害/損害を被る欠陥 | ・データの紛失, 誤ったデータの生成 ・意図しないアプリケーションの終了/ハングアップ |
| 中 | ・重, 軽のいずれにも該当しない欠陥 | ・画像の拡大機能が解除できないなど画面表示の問題 |
| 軽 | ・操作性に問題はあるが作業の継続が可能な欠陥 ・機能上は問題ないがユーザーの満足感/安心感を損なう欠陥 | ・レスポンスが悪いなど操作性の問題 ・画面のちらつき, スペルミスなど, 見た目の問題 |

テスト実施後に、検出総欠陥数、重篤欠陥の占める割合から提案の手法の適用が欠陥検出に影響するかを確認した。本来であればテスト実施者のスキルのバラツキを相殺するため、被験者の数を増やす、有識者もテストに参加させるなどの工夫が必要であるが、今回は省略した。

5.2. 実験結果

提案の手法を適用した場合と適用しない場合との欠陥検出数の違いを図(a)、同じく重篤度別検出比率を図(b)に示す。図(a)より提案の手法を適用した場合は適用しない場合と比較して、欠陥の検出数が同等もしくは増加している。図(b)より提案の手法を適用した場合は中度の欠陥が検出されているのに対して、適用しない場合に検出された欠陥はすべて軽度な欠陥である。

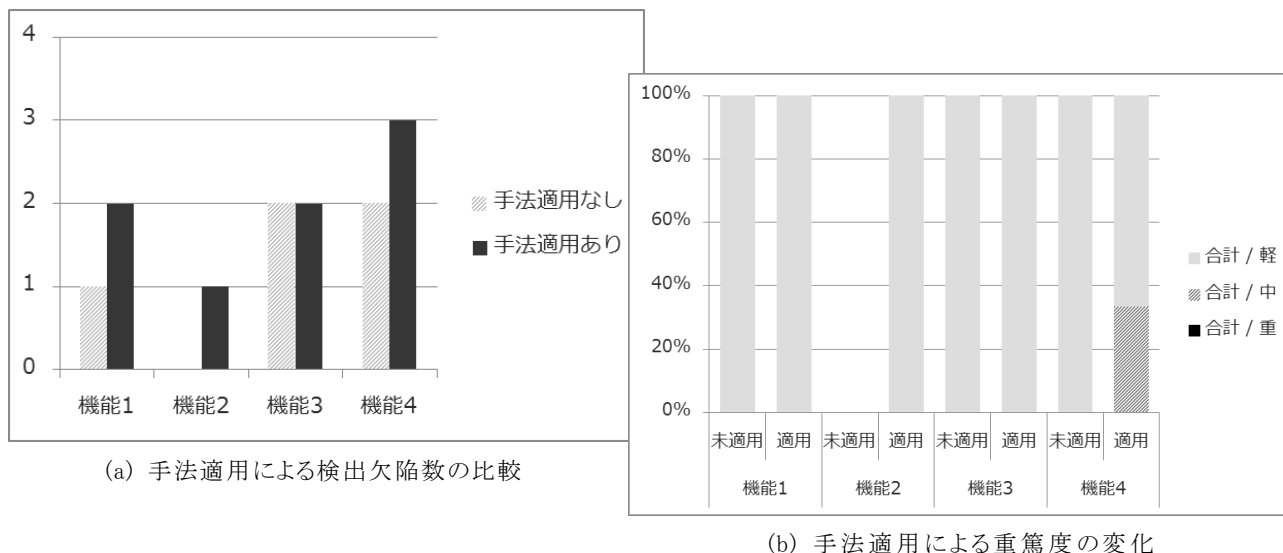


図3 提案の手法適用による欠陥検出数の比較

SOHTを見ながらのテスト実施において各チームが要した時間を表6に示す。どの機能においても提案の手法を適用した方がテスト実施の時間は長かった。

表6 テスト実施時間の比較(人/h, 下線=手法を適用)

| チーム | 機能1 | 機能2 | 機能3 | 機能4 |
|-----|------------|------------|------------|------------|
| A | 0.3 | 0.5 | <u>0.8</u> | <u>1.1</u> |
| B | <u>0.4</u> | <u>0.6</u> | 0.5 | 0.5 |

5.3. アンケート結果

アンケートの質問ごとに回答数を表7に示す。

「SOHTトレーニング法」については被験者全員が肯定的であった。「仮説の紐付け」については被験者全員が効果を感じなかった。導入の容易さに関してはややバラツキが出た。

表7 アンケートの結果

| 質問 | 回答 | | | |
|---------------------------|------------|------------|-------------|--------------|
| | 大変 そう思う | やや そう思う | あまり 思わない | まったく 思わない |
| ① 提案手法は欠陥検出に有効であると思うか。 | 1 | 3 | 0 | 0 |
| ② 提案手法を使うことでテストの幅は広がったか。 | 1 | 3 | 0 | 0 |
| ③ 提案手法はより重篤な欠陥に有効であると思うか。 | 0 | 4 | 0 | 0 |
| ④ 仮説の紐付けによって欠陥を検出できたと思うか。 | 0 | 0 | 1 | 3 |
| ⑤ 提案手法は導入が容易と思うか。 | 2 | 1 | 1 | 0 |

第5分科会（教育チーム）

提案手法に対する意見、コメントとしては次のようなものがあがった。

- ・SOHTを作成するだけでも十分トレーニングになると感じた。
- ・従来の観点だけが列挙されたリストと比較して、「状況」で括られているために実際のテストに活用しやすいと感じた。
- ・今回は仮説の関連付けにより未知の欠陥を検出することはなかったが、アイデアとしては面白いと思う。

5.4. 考察

今回の実験結果およびアンケートから、提案の手法は初級TEのテスト実施スキル向上のトレーニングとして効果があると考えた。実験結果では大きな差は出なかったが、提案の手法を適用することで中度の欠陥を検出することができた。

提案の手法を適用したテスト実施の時間が長くなったのはSOHTを活用してテストを実施したためである。初級TEはただテストケースを消化するテスト実施ではなく、欠陥を推測してテストケースにないテストも実施する上級TEのテスト実施を体験できていると考えられる。

SOHTについては「作成するだけでもトレーニングになる」、「従来の観点だけのリストより活用しやすい」とのコメントから有用であったことが伺える。

今回の実験では「仮説」の紐付けによる欠陥は検出されなかった。同一あるいは類似欠陥の推測が簡単ではないことの現れであり、初級TEのテスト実施スキルとして強化が必要であることがわかった。

6. 課題

実験結果および考察から課題を下記に述べる。

(1) サンプルの少なさ

今回の実験は被験者の人数や検出された欠陥数が少なく、統計的に効果があることは立証できていない。それを確たるものとするには被験者の人数を増やす、繰り返し実施するなど更なる実験が必要である。

(2) 同根欠陥の検出

SOHTを用いたトレーニングの実施にはテストのマネジメントが必要である。SOHTを用いたトレーニングの実施で同類仮説の紐付けによる欠陥が検出できた場合、欠陥の原因が同じである可能性が考えられる。本来、同根欠陥が存在する可能性がある場合は、検出に労した時間が無駄にならないようにテストを保留し、開発にその情報をフィードバックすることでテスト前に欠陥を除去することが望ましい。

7. まとめ

本研究では、初級TEのテスト実施スキル向上手法として、上級TEの持つ知見をSOHTとして可視化し、SOHTを用いて欠陥を検出するテスト実施トレーニング「SOHTトレーニング法」を提案した。実験により初級TEのテスト実施スキルを向上させるトレーニング手法として有用であることを確認した。

今後はサンプル数を増やし、更なる実験を行うことで提案手法の効果を評価し、改善していくことが必要である。SOHTは好評であったが、「仮説」の紐付けによる欠陥は検出されなかった。SOHTの「評価」や「更新」も含めた一連の手順を評価する必要があると考える。

第5分科会（教育チーム）

参考文献

- [1] 河野哲也, 西康晴, バグのパターンを用いたテストの提案 ～バグのナレッジマネジメント～, ソフトウェアテストシンポジウム2006 (JaSST'06), B4-3, pp.1-4, 2016
- [2] 特定非営利活動法人 ソフトウェアテスト技術振興協会 (ASTER), 一般社団法人 IT検証産業協会 (IVIA), SSFに基づくテスト技術スキルフレームワークスキル基準 -テストプロフェッショナルの戦略的育成に向けて- [Version 1.0], pp.4-6, 2012
- [3] 細川宣啓, 永田敦, 森崎修司, 川合大之, 西村英俊, 添田健太郎, 小田部健, 中谷一樹, 奥山剛, 菅野良太, 會見知史, 上野直樹, レビューオリエンテーションキットを用いた育成によるレビュー文化の醸成 -心構え, 文化, 技術観点の伝授によるレビュー人口増加-, SQiPシンポジウム2012, B1-1, pp.1-8, 2012
- [4] 中谷一樹, 原佑貴子, 上田裕之, 森崎修司, 藤田延介, 山田将貴, 原山拓実, レビュースキルを飛躍的に向上させるトレーニングの提案～レビューアのレベル定義とレベルに合わせた効率的なトレーニング～, SQiPシンポジウム2015, B3-3, pp.1-8, 2015
- [5] 中谷一樹, 上田裕之, 原佑貴子, 佐津井淑, 高橋信弘, 高橋美和子, 延原敦, 欠陥パターンと検出テクニックを用いたレビュー手法の提案とそのトレーニング教材の有効性評価の報告 -DPDT法による欠陥の検出速度及び検出精度の同時向上-, SQiP研究会, pp.1-21, 2015
- [6] 奥村有紀子, 秋山浩一, 堀田文明, 稲葉新, 松尾修, 尾崎直弘, テスト観点テンプレートを使用したテストケースの充実, SQiP研究会, pp.1-10, 2012
- [7] 奥村有紀子, 秋山浩一, 喜多義弘, 堀田文明, 川上裕司, 大島祥吾, 榎本和也, 岩尾隆弘, ソフトウェアテストにおけるバグ推測によるテスト設計手法の提案～過去バグ事例を活用した経験が浅いメンバーのバグ推測能力の向上～, SQiP研究会, pp.1-8, 2015
- [8] 光永洋, 故障事例によるテスト観点知識ベース構築とテスト設計への適用, SQiPシンポジウム 2012, B3-2, pp.1-8, 2012
- [9] 細川宣啓, 永田敦, 藤原雅明, 森崎修司, 上田裕之, 高橋功, 高橋実雄, 中谷一樹, HDR 法: 仮説駆動型レビュー手法の提案 -HDR 法の実践による生産性と品質の同時向上-, SQiPシンポジウム2013, A1-2, pp.1-8, 2013
- [10] 細川宣啓, 永田敦, 藤原雅明, 森崎修司, 中谷一樹, 井田達也, 松本達, 浅尾義則, 中島優紀, 川上隆也, 高橋功, 欠陥連鎖チャートを用いたレビュー手法の提案 -欠陥知識の有効利用によるレビュー効率の向上-, SQiPシンポジウム2014, A2-1, pp.1-8, 2014
- [11] 細川宣啓, 永田敦, 柏原一雄, 鈴木裕一郎, 田村光義, 東久保理江子, 保栖真輝, ソフトウェア欠陥予測アルゴリズム～欠陥混入メカニズムのモデリング手法を利用した欠陥予測方法の提案～, SQiP研究会, pp.1-12, 2014