

設計レビュー指標値の算出

Calculation of design review index

(株) 日立製作所 情報・通信グループ

Hitachi, Ltd. Information & Telecommunication Systems

○飯田 俊介 溝口 敦 水崎 恵一 吉田 直美

○Shunsuke Iida Atsushi Mizoguchi Keiichi Mizusaki Naomi Yoshida

Abstract In developing business application software, we conduct reviews of specifications by experts not to leave bugs that are made in a design process. However, a quantitative index that we use to evaluate reviews doesn't exist, we therefore cannot evaluate quality of design. So this paper intends to calculate a quality index of reviews by analyzing quality data of reviews and program test that are collected in one project.

1. はじめに

日立では、業務アプリケーションを開発するに当たって、設計工程で作りこまれた不良（以下、設計不良と呼ぶ）を次工程以降で発生させないようにするために、有識者による設計書のレビューを実施している。これは、次工程以降で設計不良が発生すると、設計工程まで立ち戻って原因を調査するため、大幅な手戻りが発生してしまい、納期遅延につながってしまう恐れがあるからである。したがって、設計段階で十分にレビューを行い、不良を早期に摘出することが、開発を進めていくにあたって重要となる。

しかし、このレビューの十分性評価に用いる定量的な指標が存在しないため、設計品質について定性的評価しか行えない現状である。実際に、あるプロジェクトのテスト工程では、設計レビューが十分に行なわれていないことにより、設計不良が多数発生している。このようなことを防止するために、設計レビューに対して定量的な指標を立てて、レビューの十分性を評価することが重要となる。

そこで本論文では、1つのプロジェクト（以降、Aプロジェクトと呼ぶ）を分析対象として、レビューの議事録に記載されている詳細設計書の頁数や実施時間、指摘件数といったレビューに関する品質データと、テスト工程における品質データを分析し、設計レビューの品質指標値を導き出すことを目的とする。

2. データの集計

本章では、分析対象としたプロジェクトの概要と、分析するデータの概要、及びその集計方法について説明する。

2.1 分析対象プロジェクト

本論文では、金融系システムの業務アプリケーション開発（Aプロジェクト）から得られたデータを分析対象とした。分析対象としたAプロジェクトの規模や開発手法などの特徴を表1に示す。

(株) 日立製作所 情報・通信グループ 品質保証本部 金融システム品質保証部
Hitachi, Ltd., Information & Telecommunication Systems Quality Assurance & Inspection
Department of Financial Information Systems Division

川崎市幸区鹿島田 890 (日立システムプラザ新川崎) 〒212-8567 Tel:044-549-1518
890 Kashimada, Saiwai, Kawasaki, Kanagawa, 212-8567 Japan

表1. Aプロジェクトの特徴

NO	特徴	概要
1	開発言語	J a v a
2	規模	500 [KLOC]
3	開発手法	ウォーターフォールモデルに沿って開発を進める。なお、日立における開発の考え方である、V字モデルの概要を図1に示す。
4	コーディング	ツールによる自動生成は行わない。

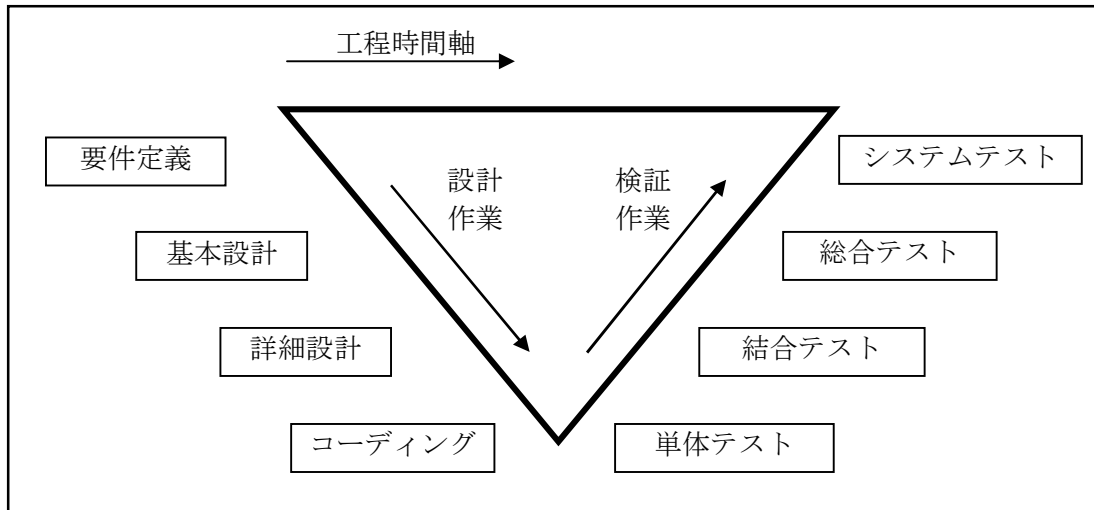


図1. V字モデル

2.2 分析対象データ

詳細設計書のレビューにおける品質指標値として、レビューでの指摘数と実施時間を扱う。テスト工程で発生した設計不良¹は、詳細設計時に作りこまれたものである。詳細設計書のレビューで抽出できなかった不良は、その後のテスト工程で抽出される。つまり、レビューが十分に行なわれていなければ、テスト工程で多くの設計不良が抽出されるが、レビューが十分に行なわれ、不良を抽出することが出来れば、テスト工程では設計不良は抽出されず、製造不良のみとなる。このことから、レビューでの指摘数や実施時間と、テスト工程における設計不良とは相関があると考えられる。そこで本論文では、レビューを実施した結果とテスト結果を分析し、品質指標値を導く。表2に分析対象データの概要を示す。

表2. 分析対象データ

NO	分析対象データ	概要
1	指摘密度	レビューにおける詳細設計書1ページあたりの指摘件数。誤字脱字に関する指摘は含まない。
2	時間密度	詳細設計書1ページあたりのレビュー実施時間。
3	不良密度	単体テスト、結合テストにおける不良のうち設計不良に起因する不良数。本来はレビューで抽出すべき不良が対象。機能量で正規化するために対象プログラムのドキュメント頁数を用いて不良密度を算出。

¹ 設計不良：テスト工程で抽出された不良のうち、詳細設計書とプログラムの両方を改修する必要がある不良を示す。本来はレビューで抽出すべき不良が対象。

指摘密度、時間密度はレビュー議事録から、不良密度はバグ票から、収集した。また、Aプロジェクトでは、ユースケース（以下、UCと略す）²ごとに一人の担当者が設計し、レビューもユースケースごとに実施していたため、データはユースケース単位に集計した。

（1）データの前提条件

品質指標値を求めるにあたって、Aプロジェクトでの詳細設計書の形式や、レビューの実施形式について述べる。以降のデータは、全て表3の条件のもとで得られたデータである。表3にデータ分析で扱う詳細設計書、レビューの実施形式に関する前提条件を示す。

表3. 前提条件

NO	項目	内容
1	詳細設計書	以下の作成すべき全てのドキュメントが揃っていること。 <ul style="list-style-type: none"> ・バウンダリコンポーネント³仕様書 ・コントロールコンポーネント⁴仕様書 ・画面項目仕様書 ・シーケンス仕様書 ・SQL 仕様書 ・テーブル設定項目仕様書 ・電文設定仕様書 ・帳票設計書
2	レビュー実施形式	ウォークスルー形式であること。
3		レビューアに対する資料の事前配布は行わない。
4		レビューアとして各コンポーネントのリーダー、サブリーダーが出席していること。なおリーダーとサブリーダーは、類似システム開発経験を持つ。

（2）集計データの絞込み

全ユースケースのうち、データ不足や、異常値を除くため、以下の条件に合致するユースケースを分析対象から除外する。

- ① 詳細設計書レビューの議事録そのものがない、あるいはレビュー時間の表記がないなど、レビュー結果を数値化できないもの。
- ② 開発対象規模が小さいもの（詳細設計書が100頁未満）。開発規模が小さいので、規模で正規化した際に1件の不良でも大きな不良密度となってしまうため、異常値として扱う。条件①、②で分析対象を絞った結果、8UCを分析の対象とした。

（3）説明変数と目的変数の算出

分析対象となるUCから詳細設計書レビューにおける指摘密度、時間密度と、テストでの不良密度を算出したものを表4に示す。なお、指摘密度と時間密度を回帰分析における説明変数（変数：x）、不良密度は目的変数（変数：y）と記述する。

² ユースケース（UC）：エンドユーザにとって1つのサービスの開始画面から終了画面までの一連の流れを1UCとする。

³ バウンダリコンポーネント：画面から渡された値をコントロールコンポーネントで使用するために編集したり、コントロールコンポーネントから渡された値を画面に出力するために編集したりする処理を行う。

⁴ コントロールコンポーネント：主に業務ロジック処理を行う。

表4. 分析対象UCのレビュー密度と不良密度

NO	ユースケース名	説明変数 (変数: x)		目的変数 (変数: y)
		指摘密度[件数/頁]	時間密度[h/頁]	不良密度[件数/頁]
1	ユースケースA	0.25	0.07	0.004
2	ユースケースB	0.24	0.04	0.01
3	ユースケースC	0.22	0.03	0.02
4	ユースケースD	0.15	0.03	0.03
5	ユースケースE	0.04	0.02	0.02
6	ユースケースF	0.13	0.03	0.01
7	ユースケースG	0.10	0.04	0.02
8	ユースケースH	0.04	0.01	0.02

3. 分析及び結果

3.1 分析手法

本論文では、算出する品質指標値を前章で述べた2種類（指摘密度、時間密度）のレビュー密度とする。それぞれのレビュー密度の最適値は共に以下の手順で求める。

- ① 指摘密度、時間密度と不良密度の相関係数を求める。
- ② 両者に相関関係が認められることを確認し、回帰直線を求める。
- ③ ②で得られた回帰直線から不良数が0となるそれぞれのレビュー密度を求める。

③で別々に得られた指摘密度、時間密度が、求めるべきレビューの品質指標値とする。

3.2 相関係数の算出

表3で集計したデータを元に、2種類のレビュー密度とテストで発生した不良密度との相関分析を行う。なお、相関係数Rと相関の強さは一般的に表5のような関係があるといわれている。

表5. 相関係数と相関の強さ

R	相関の強さ
0~0.2	ほとんど相関がない
0.2~0.4	弱い相関がある
0.4~0.7	中程度の相関がある
0.7~1.0	強い相関がある

表4を相関分析した結果から得られるレビュー密度と不良密度との相関係数Rは、表6のようになる。

表6. 相関分析結果

NO	説明変数: x	目的変数: y	相関係数: R
1	指摘密度	不良密度	-0.58
2	時間密度		-0.70

表5と、表6の得られた相関係数を比較した結果、次のことが結論づけられる。

- 1) 指摘密度、不良密度の間には「中程度の相関」がある。
- 2) 時間密度、不良密度の間には「強い相関」がある。
- 3) とともに負の相関があり、指摘件数が多いほど、また、レビュー時間が長いほどテスト工程

まで残る設計不良が減る。

3.3 回帰直線の算出

3.2節で2つのレビュー密度と不良密度に相関があることが示された。次にそれぞれの回帰分析を行い、回帰直線を求める。

(1) 指摘密度と不良密度との散布図と回帰直線を図2に示す。

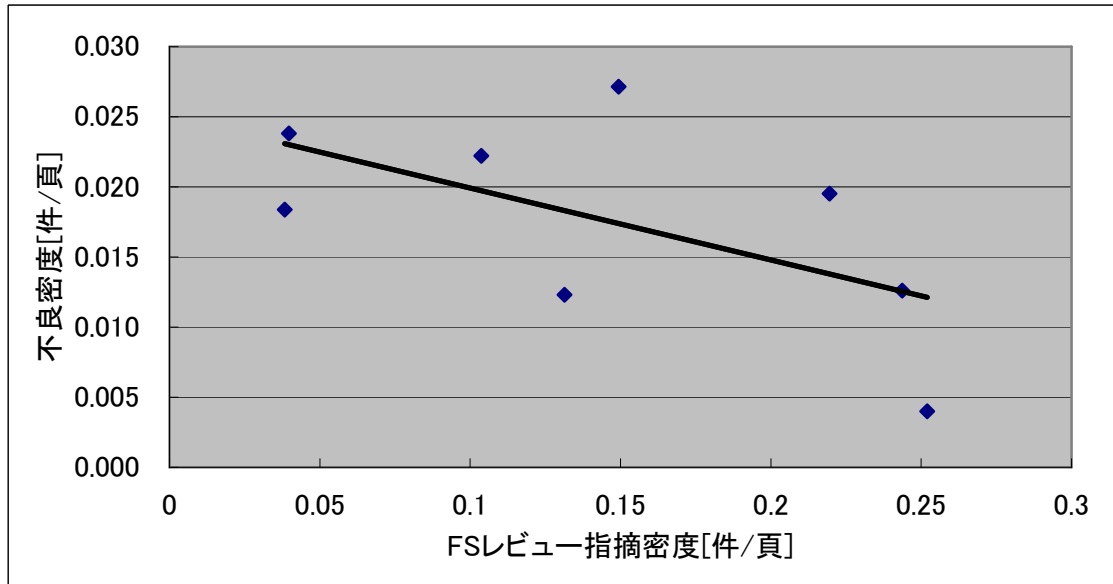


図2. 指摘密度、不良密度の散布図および近似曲線

得られた回帰直線は以下の1次方程式で表される。

$$y = -0.051x + 0.025 \cdots(1)$$

(2) 時間密度と不良密度との散布図と回帰直線を図3に示す。

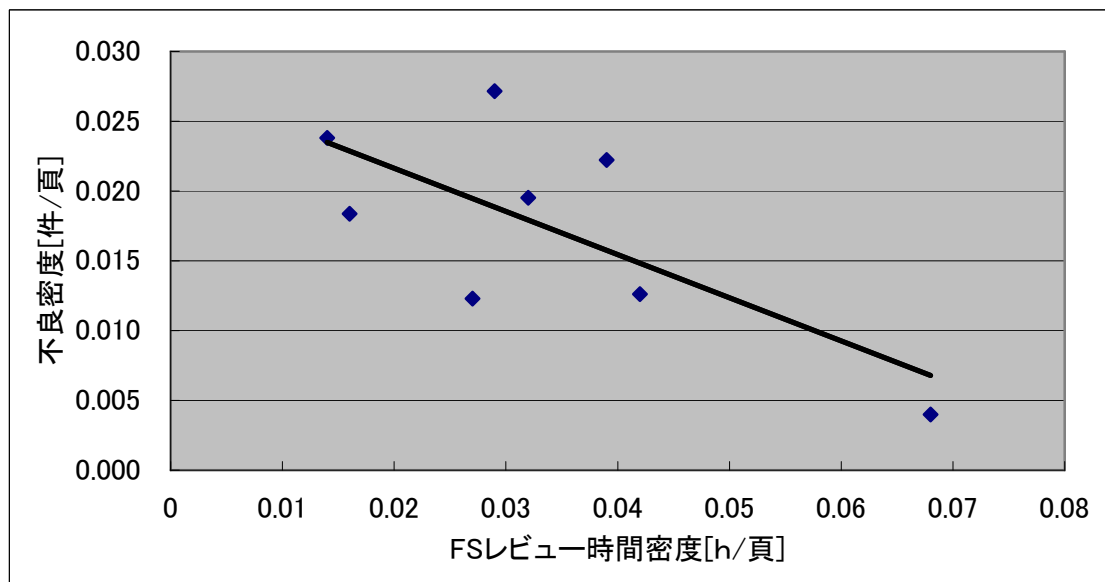


図3. 時間密度、不良密度の散布図および近似曲線

得られた回帰直線は以下の1次方程式で表される。

$$y = -0.309x + 0.028 \cdots(2)$$

3. 4 レビュー品質指標値の算出

3. 2節、3. 3節では、2つのレビュー密度と不良密度との間にはともに負の相関関係があることを示し、それぞれの2変数を表す1次方程式を求めた。この結果からレビューにおける品質指標値を求めることを考える。求めるべき指標値は不良密度を限りなく0にするためのレビュー密度であるので(1)式、(2)式にそれぞれ $y=0$ を代入することで求めることが出来る。結果は以下の通りになる。

指摘密度：0.49[件/頁]

時間密度：5.4 [分/頁]

以上から指摘件数では10頁あたり約5件の指摘、レビューにかかる時間では10頁あたり約54分とすることで、テスト工程において、設計不良を限りなく0件に近づけることが出来るという分析結果が得られた。

4. まとめと今後の課題

本論文では、詳細設計書についての品質指標値を導き出すために、レビューでの指摘密度とテスト工程での設計不良に起因する不良を纏め、相関分析を行った。その結果導き出された相関係数を基に回帰分析を行い、レビューの品質指標値となる指摘密度、レビュー時間を算出することができた。

今回導き出した品質指標値は、1つのプロジェクト固有の指標値にしかないが、今後この考え方をを用いて複数のプロジェクトのデータを分析し、得られた指標値を用いることで、以下の2点が期待される。

(1) 求めた指標値をレビュー時の目標値として用いることで、レビューの十分性を評価することが出来る。

(2) これまではレビューや作業のプロセスのみを見て、全体の設計品質を評価していたが、レビュー時の指摘件数や実施時間と、指標値とを比較することで、設計品質が確保されているかどうかを評価することが可能となる。

今後の課題としては、分析対象プロジェクトを増やして、データの信頼性を上げていくこと、レビュー結果が人的要因に左右されない仕組みを作ることが挙げられる。後者については、レビューアによるレビューの質の差や、設計者による設計書の品質の差という、2つの人的要因を事前に排除してレビューを実施することが必要と考える。そのために、標準的なレビューチェックリストや、各種設計書の作成基準を策定することが今後の取り組むべき課題である。