

ソフトウェア信頼性技術(2)

5K-4

-遅延S字形NHPPモデルの安定化-

樋口 潔¹ 阿部 勝徳² 上村 松男¹ 寺本 雅則¹¹日本電気(株) ²日本電気マイコンテクノロジー(株)

1 まえがき

ソフトウェア信頼度成長モデルを用いてエラー発見過程を予測する場合、予測が正しいかどうかということは重要な問題である。予測が正しいければ、それ以後からリリース時までのエラー発見過程を正確につかむことができる。すなわち、検査の計画が立てやすくなるわけである。一般には、発見エラーの実績データが多いほど精度の良い予測ができると考えられている。

またソフトウェアは、リリース後にちがう環境で使われる。そのため、検査段階とリリース後ではエラー発見過程も異なっている。

本稿では、検査段階のエラー発見過程を正確に予測するために必要な実績データ収集期間について、まず考察する。また、リリース後におけるソフトウェア信頼度成長モデルの適用についても議論する。

2 遅延S字形NHPPモデル

ソフトウェア信頼度成長モデルとして、遅延S字形NHPPモデル[1]を用いた。このモデルは、パラメータの数が少なく、そのおのおのの意味が明確であるという利点がある。さらに、このモデルによって信頼度も定義できる。

このモデルでは、 t 時点までに発見されたエラーの数 $M(t)$ を次のように推定する。

$$M(t) = a \{1 - (1 + bt) \exp[-bt]\} \quad (1)$$

ここで、 a は潜在総エラー数、 b は期待発見率(発見率の上限)をあらわす。

また、 t 時点の信頼度 $R(x|t)$ は次のように与えられる。

$$R(x|t) = \exp[-\{M(t+x) - M(t)\}] \quad (2)$$

ここで、 x は信頼度区間である。今回の分析で適用したソフトウェアが、すべて発見エラー数を1週間ごとに計測していたため、 x を7日間とした。

3 分析データ

この分析では、以下に示す2種類のソフトウェアを適用した。

CS: 通信ソフトウェア

このソフトウェアは、デジタル交換機用で、リアルタイム性のきわめて高いシステムである。また、交換システム特有の厳しい運転環境を考慮した、信頼性が要求された。システムは、V1という初版から始めて逐次機能を拡張している。V1の全体規模は、50KLである。V2は、V1の機能を拡張したものであり、全体の規模は新規3KL、流用47KL、改造10KLの計60KLである。いずれも Real Time Cと呼ばれるC言語の一種で記述されている。

CC: Cコンパイラ

これは、リアルタイムOS用Cコンパイラである。バ

ージョンは、V1.0から始まって、V1.8までの9つがある。ただし、各々のバージョンアップは、エラーの修正を反映するためのものであった。エラーデータは、V1.0をリリースする以前の検査段階以降、すべてのものが収集されている。CCの全体の規模は、24KLで、そのうち新規14KL、流用9.4KL、改造0.6KLである。システムはC言語で記述されている。

4 検査段階におけるモデルの安定

“モデルが安定する”とは、最後まで正しい予測をするモデルが得られ、それ以後の実績発見エラー数を入力しても a の値がほとんど変わらなくなった状態(リリース時点の95%信頼区間から a の値がはずれなくなった状態)をいうことにする。つまり、この時点でかなり正確な見積りができる。どこまで検査を進めれば、モデルが安定するのか、CSのV1, V2を用いて検討する。

次の図1は、V1の潜在総エラー数の推定値 a と95%信頼区間の推移を表す。(月ごとに、その時点までに得られた発見エラーの時系列データによって a, b を推定した。)この図から、ある時点 t_k までは a の推定値が激しく変化し、それ以後は、ほぼ同じ値をとることがわかる。(リリース時点の95%信頼区間に、 a の値が入り続ける。)V2でも、このような傾向が見られた。この t_k がどの時点なのかを検討することで、モデルの安定する時点がわかる。

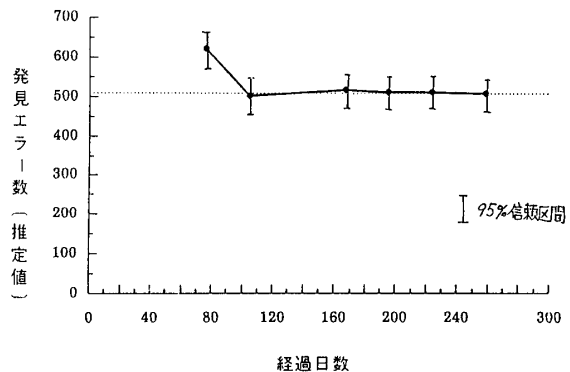


図1 リリース時点の発見エラー数推定値と95%信頼区間の変化(CS V1)

まず、検査開始から、リリース時点までを100とする。このとき安定する時点 t_k は、V1が40.5時点、V2が57.4時点であった。この結果からは、安定する時点の規則性が見えない。これは、V1とV2のリリース時の信頼度が異なるためだと考えられる。(V1は83.7%、V2は、60.0%だった。)

そこで次に、V1, V2とも80.0%でリリースしたと仮定して t_k を計測する。検査開始から、信頼度80.0%の時点までを100とする。このとき、安定する時点 t_k は、V1が41.8

時点、V2が49.1時点であった。ここでは、安定する時点が検査段階の中程であることがわかる。また、 t_k 時点の信頼度は、V1が0.0007%、V2が2.97%であり、かなり信頼度が低い時点で安定していることもわかった。

以上のことから、検査工程が40~50%進んだ時点（または信頼度が、数%の時点）で、モデルが安定し、検査要員の配置、テストのしかたなどを考え直すことができるといえた。

しかし、これはCSのみから得られたものなので、今後のデータの積み上げが必要であることは言うまでもない。

5 リリース後におけるモデルの安定

ここでリリースとは、社内向け限定リリースのことをいう。CCは、リリース後のエラー発見過程が得られている。バージョンアップは、エラーの修正を反映するためだといってもよいので、リリースとエラーの増え方の関係がよく見える。これを用いて3章で述べたモデルの安定を、リリース後についても考察する。

図2は、CCの発見エラー数実績（SOREMによる出力）である。（SOREMについては、[2]のソフトウェア信頼度技術（1）で述べた。）全体的に、おおむねS字のカーブをなしている。しかし、V1.0、V1.5、V1.7の各リリース時に、小さいS字の切れ目が見られる。原因は、以下のようなことがとしては、V1.0リリース時は、このソフトが初めてリリースされた時であったからである。V1.5の時は、このソフトウェアがVLSIチップに組み込まれ、大量の新しいユーザがついたからである。またV1.7の時は、あるバグが修正され、このことによって、それまで動かなかった部分が動くようになったからである。このような利用環境の変化により、突然エラー発見数が増えることがわかっているが、この図はそれをよく表している。

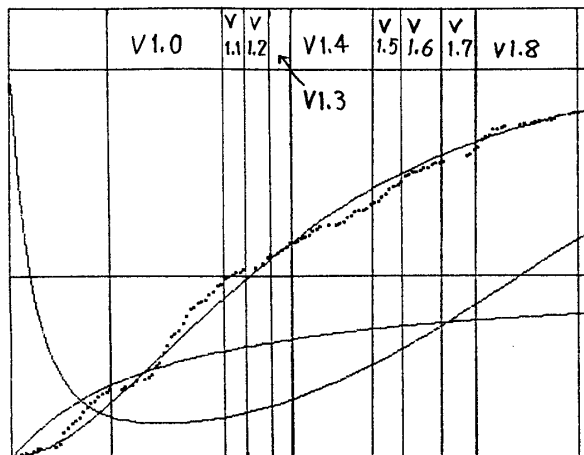


図2 CCの発見エラー数とバージョン

モデルが安定しているかどうかは、図3を見るとわかる。V1.2リリース以前は、潜在総エラー数の推定値が、激しく変化している。その後、V1.5リリース前までは、潜在総エラー数の推定値が安定した。しかし、その後V1.8まで潜在総エラー数の推定値が増え続けている。モデルの安定に関しても、利用環境の変化に対する影響が大きい。前述のV1.5とV1.7での特徴をよく表している。ただし、V1.0の影響はわからない。これは、V1.0以前の検査段階データを細かくとれば、わかるであろう。

以上のように、潜在総エラー数の推定値はリリースする度に大きくなる傾向がある。リリースすると、テスト

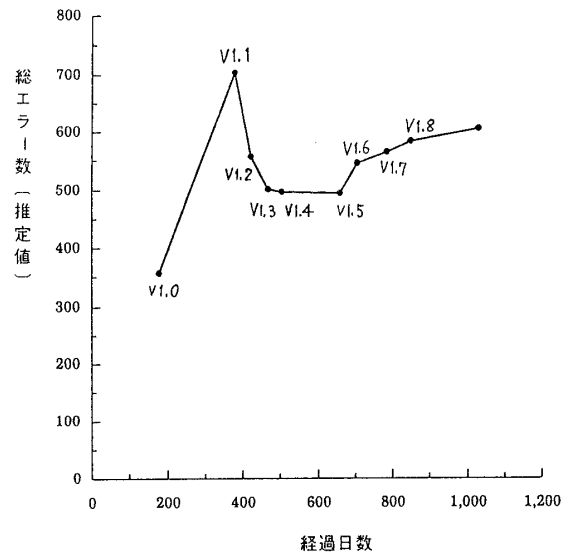


図3 CCの潜在総エラー数推定値の変化

の十分されていなかった部分がユーザに使われたり、エラーを修正したために使えなかった部分が使えようになつたりする。そのために潜在総エラー数の推定値が、増加するのである。つまり、潜在総エラー数の推定値は、テストの網羅度とも関係する値だと言える。

6 あとがき

本稿では、遅延S字形NHPPモデルの安定性について述べた。これによってモデルの安定する時点が、検査を40%~50%終了したところだということがわかった。この結果は、他のプロジェクトでも適用できると考えられる。また、リリース後は環境も変化するため、いちがいにモデルが安定する時点を決められないことがわかった。

今後は、いろいろな尺度（規模、発見率、テストケースの選び方など）とモデルの安定する時点との関係を調べ、安定時点をもさらに正確に予測できるようにする必要がある。また、モデルの安定時点を早める方法を検討する必要もある。

参考文献

- [1] 山田茂、尾崎俊治：“ソフトウェアエラー発見過程に関する信頼性モデル”、情報処理学会論文誌 Vol. 24, No. 3, 1983
- [2] 上村、青木、藤野：“ソフトウェア信頼度技術—信頼度メトリックス—”、情報処理学会第35回全国大会論文集（Ⅱ）、1987