

TIPSH アルゴリズムに基づいたウェーブレット縮小推定による ソフトウェア信頼性評価

広島大学大学院工学研究科 肖 霄, 土肥 正

Xiao Xiao and Tadashi Dohi
Graduate School of Engineering
Hiroshima University, Japan

1 はじめに

ソフトウェア信頼性モデル (Software Reliability Model; SRM) は, テスト段階で観測されたフォールト検出に関する各種統計情報に基づいて, ソフトウェア障害が発生する現象を記述する数学モデルであり, ソフトウェア信頼性を定量的に評価するために利用される. 一般に, ソフトウェアフォールト検出過程を非同次ポアソン過程 (Non-Homogeneous Poisson Process; NHPP) に従うと仮定することが多く, NHPP に基づいた SRM はソフトウェアの信頼性評価を実施する上で最も広範に適用されている [6]. これまでに提案された NHPP に基づいた SRM の多くはパラメトリックモデルであった [8]. パラメトリックモデルでは期待累積フォールト数を表す平均値関数に滑らかな関数形を仮定するため, 日々のテストにおいて変動する検出フォールト数の詳細な振舞いを原理的に記述できないことが弱点として挙げられる. よって, モデリングにおけるソフトウェアデバッグに関するシナリオを必要としないノンパラメトリックモデルの重要性が増してきている.

近年, 計数過程のノンパラメトリック推定を行うウェーブレット縮小推定 (Wavelet Shrinkage Estimation; WSE) が急速に整備されてきた [1]. Xiao and Dohi [9] では WSE を NHPP に基づいた SRM のノンパラメトリック推定法として初めてソフトウェア信頼性分野に導入し, ソフトウェア信頼性評価における WSE の有効性について検証した. Xiao and Dohi [9] は特に, データ変換を伴う WSE に着目し, 従来のパラメトリック推定法である最尤法 (Maximum Likelihood Estimation; MLE) や最小二乗法 (Least Squares Estimation; LSE) に基づいた評価よりも単位テスト時間当たりの検出フォールト数を高い精度で推定できることを示した. 一方, Kolaczyk [7] はポアソン過程の強度関数を WSE で推定する際にデータ変換が原因で oversmoothing などの問題点が生じやすいことを指摘し, 自らが導出した閾値を使用することによってデータ変換を行わない WSE を提案した. その実装のために TIPSH (Translation-Invariant Poisson Smoothing using Haar Wavelets) アルゴリズムを開発し, シミュレーション実験を通して TIPSH アルゴリズムの有効性を示した. 本稿では, TIPSH アルゴリズムを実際のテスト工程で観測されたソフトウェアフォールトデータの解析に適用し, データ変換を行わない WSE について考察するとともに, 文献 [9] における評価結果との比較を行う. 以降では 2 つの手法を区別するために, データ変換を伴う WSE を DT-WSE, データ変換を伴わない WSE を NDT-WSE によって表記する.

2 ウェーブレット縮小推定

ウェーブレット縮小推定の基本原理は, まずノイズを含んだ観測データをウェーブレット展開した後, ウェーブレット係数を抽出し, 閾値法 (thresholding) によってノイズ除去 (denoise) を行う. そしてディノイズされた係数をウェーブレット展開式に代入することで, 観測データの背後に隠さ

れた真の関数を推定するというものである。Donoho and Johnstone [3, 4] による一連の研究では、ウェーブレット推定の統計的合理性に関する重要な性質が証明されている。

2.1 ハールウェーブレット変換

既に提案されているウェーブレット系として、ハールウェーブレット系 (Haar Wavelets), シヤノンウェーブレット系 (Shannon Wavelets), ルマリエ-メイエウェーブレット系 (Lemarié-Meyer Wavelets) 等があるが、本研究では離散データの処理に適しているハールウェーブレット系に着目する。ハールのスケーリング関数 (scaling function) とウェーブレット関数 (wavelet function) は、それぞれ、

$$\phi(i) = \begin{cases} 1 & (0 \leq i < 1) \\ 0 & (\text{otherwise}), \end{cases} \quad (1)$$

$$\psi(i) = \begin{cases} 1 & (0 \leq i < \frac{1}{2}) \\ -1 & (\frac{1}{2} \leq i < 1) \\ 0 & (\text{otherwise}) \end{cases} \quad (2)$$

によって定義され、拡大縮小のパラメータ j と平行移動のパラメータ k を加えると、ハールのファザーウェーブレット (father wavelet) とマザーウェーブレット (mother wavelet) は

$$\phi_{j,k}(i) = 2^{j/2} \phi(2^j i - k), \quad (3)$$

$$\psi_{j,k}(i) = 2^{j/2} \psi(2^j i - k) \quad (4)$$

のように生成される。

解析対象関数 λ_i ($i = 1, 2, \dots, n$) のウェーブレット展開は以下のように与えられる。

$$\lambda_i = \sum_{k=0}^{2^{j_0}-1} c_{j_0,k} \phi_{j_0,k}(i) + \sum_{j=j_0}^{\infty} \sum_{k=0}^{2^j-1} d_{j,k} \psi_{j,k}(i). \quad (5)$$

ここで、

$$c_{j_0,k} = \sum_{i=1}^n \lambda_i \phi_{j_0,k}(i), \quad (6)$$

$$d_{j,k} = \sum_{i=1}^n \lambda_i \psi_{j,k}(i) \quad (7)$$

は、それぞれスケーリング係数 (scaling coefficient) 及びウェーブレット係数 (wavelet coefficient) と呼ばれ、 j_0 (≥ 0) は初期解像度レベルを表す。

スケーリング係数とウェーブレット係数は基底関数を通して解析対象関数の各微小区間における情報を抽出し、それらを足し合わせることで解析対象関数の滑らかさと細かさを表現している。従って、解像度レベルを表す拡大縮小のパラメータ j は、観測データから有用な情報を抽出する上で重要な役割を担っている。 j が小さい時、基底関数の幅が小さくなるため、時間軸上の変化を敏感に捉えることができるが、参照するデータ量が少なくなるため、各微小区間の情報を抽出する

時の精度が悪くなる。逆に j が大きい時、各微小区間における精度が向上する反面、時間軸を小刻みに表現できないため、時間に伴う変化に追従しにくくなる。従って、式 (5) のように、滑らかさに関しては初期解像度レベル j_0 のスケーリング係数によって制御し、細かさに関しては初期解像度レベル j_0 から無限大までのウェーブレット係数をすべて利用することでこのトレードオフ関係を解決し、滑らかでかつ細かい情報を含めた推定結果を得ることができる。実際の観測で得られるデータは有限個なので、解像度レベルは無限になることはなく、本稿では整数 J で最高解像度レベルを表すことにする。このように、ウェーブレット基底関数 (3), (4) は時間軸上で平行移動や拡大縮小するパラメータを持つため、解析対象関数の大域的な特性と局所的に細かい挙動を同時に捉えることができる。関数 λ_i から係数 $(c_{j_0,k}, d_{j,k})$ への変換はハールウェーブレット変換 (Haar Wavelet Transform; HWT) と呼ばれる。

2.2 閾値法

閾値法では、ウェーブレット係数 $d_{j,k}$ を閾値と比較し、閾値よりも絶対値が小さい係数を 0 で置き換える。この置き換えによって小さな変動を無視することができ、データから大きなノイズを取り除くことが可能となる。一般に、ハード閾値法 (hard thresholding)

$$\delta_\tau(u) = u1_{|u|>\tau} \quad (8)$$

とソフト閾値法 (soft thresholding)

$$\delta_\tau(u) = \text{sgn}(u)(|u| - \tau)_+ \quad (9)$$

がある。ここで、 $\tau (\geq 0)$ は閾値、 1_A は事象 A の指示関数、 $\text{sgn}(u)$ は符号関数、 $(u)_+ = \max(0, u)$ である。

3 ウェーブレット縮小推定によるソフトウェア信頼性評価

3.1 離散型 NHPP モデル

ここでは単位テスト時間当たりの検出フォールト数の振舞いに着目するため、一般性を失うことなく、離散型 NHPP モデルについて述べる [10]。

今、時間区間を適当なスケールで離散化した時刻列 $i = 0, 1, 2, \dots$ を考える。テスト時刻 i において観測されるソフトウェアフォールト数を Y_i 、その累積値を $N_i = \sum_{k=0}^i Y_k$ で表し、 $Y_0 = N_0 = 0$ とする。このとき、確率過程 $\{N_i : i = 0, 1, 2, \dots\}$ が離散型 NHPP であるとは、その確率関数が

$$\Pr\{N_i = m\} = \frac{\{\Lambda_i\}^m}{m!} \exp\{-\Lambda_i\}, \quad m = 0, 1, 2, \dots \quad (10)$$

によって与えられることをいう。ここで、 $E[N_i] = \Lambda_i$ は離散型 NHPP の平均値関数と呼ばれ、時刻 i までに検出される期待累積フォールト数を表す。また、強度関数 $\lambda_i = \Lambda_i - \Lambda_{i-1}$ は $i (\geq 1)$ 日目のテストにおいて新たに検出された期待フォールト数であり、 $E[Y_i] = \lambda_i$ が成立つ。これより、離散型 NHPP モデルは平均値関数 (強度関数) によって特徴付けられ、離散関数 $\Lambda_i (\lambda_i)$ の形状によって多様なフォールト検出パターンを表現することが可能である。

ウェーブレット縮小推定を適用するために、時間区間 $[0, T]$ を $n (= 2^J; J = 1, 2, \dots)$ 等分することを考える。説明の便宜上、 $y_i = x_i - x_{i-1}$ ($i = 1, 2, \dots, n$) は強度 λ_i を持つポアソン確率変数 Y_i の実現値とする。即ち、

$$Y_i = \lambda_i + \eta_i, \quad i = 1, 2, \dots, n. \quad (11)$$

ただし, η_i はポアソン型白色雑音を表す確率変数であり, x_i ($i = 0, 1, 2, \dots, n$) は各時間区間までに検出される累積フォールト数を表す.

3.2 ディノイズの原理

一般に, 閾値法の正当性はウェーブレット係数 $d_{j,k}$ が正規分布に従うことで保証される. 文献 [5] より, 標準正規分布に従う互いに独立な n 個の確率変数 Z_i ($i = 1, 2, \dots, n$) に関して, 次のような漸近的な性質が成り立つことが分かっている.

$$\Pr\left(\max_{1 \leq i \leq n} |Z_i| \leq \sqrt{2 \log n}\right) \rightarrow 1 \text{ as } n \rightarrow \infty. \quad (12)$$

これは, n が大きくなるにつれ, ほとんどの確率変数 Z_i は $\sqrt{2 \log n}$ より小さくなることを意味する. 観測データ y_i が標準正規分布に従うなら, ウェーブレット係数 $d_{j,k}$ も標準正規分布に従うとみなすことができるので, ほとんどのウェーブレット係数 $d_{j,k}$ は $\sqrt{2 \log n}$ よりも小さいことが分かる. Xiao and Dohi [9] は閾値法を施す前に予めデータ変換を行い, フォールトデータを近似的に正規分布に従うデータに変換した上で, この性質から得られた universal threshold, $\tau = \sqrt{2 \log n}$ を使用した.

しかし, データの前処理を行うと, ほとんどのウェーブレット係数 $d_{j,k}$ が閾値より小さい値になり, 本来データの特徴をきちんと捉えているウェーブレット係数であったにもかかわらずその値が 0 に置き換えられて, 結果的にノイズではない部分も除去されてしまうことがよくある. この問題点は oversmoothing と呼ばれる. そこで Kolaczyk [7] は, データの前処理を行わず, ポアソン性を持つデータのノイズ除去に適切な閾値 level-dependent thresholds

$$\tau_j = 2^{-\frac{J-j+2}{2}} \left\{ 2 \log(2^j) + \sqrt{4 \log^2(2^j) + 8 \lambda_0 \log(2^j) 2^{(J-j)}} \right\} \quad (13)$$

を提案した. ただし, λ_0 は入力のサンプル平均である.

式 (13) の導出過程を以下に示す. 観測データ y_i がポアソン分布に従うとき, ウェーブレット係数 $d_{j,k}$ はパラメータ α, β を持つ 2 つのポアソン分布の差に比例した分布に従う. 即ち,

$$d_{j,k} \sim 2^{-(J-j)/2} \{Po(\alpha) - Po(\beta)\}. \quad (14)$$

ここで, α, β は j, k によって区切られた時間区間上の強度を表しており, 帰無仮説 $H'_0: \lambda_i = \lambda_0$ の下では $\alpha = \beta = 2^{J-j} \lambda_0 / n$ となる. 確率変数 $P_1 \sim Po(\alpha)$, $P_2 \sim Po(\beta)$, $X \sim \chi_{(\nu)}^2(2\beta)$ に対して,

$$\Pr(P_1 - P_2 \geq \frac{\nu}{2}) = \Pr(X < 2\alpha) \quad (15)$$

が成り立つことから,

$$\Pr\left(2^{(J-j)/2} d_{j,k} \geq \frac{\nu}{2}\right) = \Pr\left(\chi_{(\nu)}^2(2\beta) < 2\alpha\right) \quad (16)$$

を得ることができる. よって, 解像度レベル j ごとに最大となる $d_{j,k}$ がある閾値 τ_{J-j} より小さい確率は

$$\begin{aligned} \Pr\left(\max_{0 \leq k \leq n_{J-j}-1} |d_{j,k}| \leq \tau_{J-j}\right) &= \left[1 - 2 \Pr\left(d_{j,k} > \tau_{J-j}\right)\right]^{n_{J-j}} \\ &= \left[1 - 2 \Pr\left(\chi_{(2^{(J-j)/2+1} \tau_{J-j})}^2(2\beta) < 2\alpha\right)\right]^{n_{J-j}} \\ &= \left[1 - 2 \Pr\left(\chi_{(2^{(J-j)/2+1} \tau_{J-j})}^2\left(\frac{2^{J-j} \lambda_0}{n}\right) < \frac{2^{J-j} \lambda_0}{n}\right)\right]^{n_{J-j}} \end{aligned} \quad (17)$$

となる。ただし、 $n_{J-j} = 2^j$ である。さらに、これを

$$\Pr \left(\max_{1 \leq i \leq n} |Z_i| \leq \sqrt{2 \log n} \right) = \left[1 - 2 \Pr \left(Z > \sqrt{2 \log n} \right) \right]^n \quad (18)$$

と比べると、以下の式を得ることができる。

$$\Pr \left(\chi_{(2^{(J-j)/2+1} \tau_{J-j})}^2 \left(\frac{2^{J-j} \lambda_0}{n} \right) < \frac{2^{J-j} \lambda_0}{n} \right) \approx \Pr \left(Z > \sqrt{2 \log(n_{J-j})} \right). \quad (19)$$

この式を近似的に解くことによって、式 (13) の level-dependent thresholds τ_j が得られる。この閾値を用いることで、ポアソン性を持つソフトウェア信頼性データに対して正規変換を施す必要はなく、直接ウェーブレット変換を行うことで真の関数を推定することが可能となる。

3.3 Translation-Invariant Poisson Smoothing using Haar Wavelets

ハールウェーブレットを用いた解析では、その固有性質から推定結果には階段状のノイズが含まれることが多い。本研究ではこの問題を解決するために、Coifman and Donoho [2] が提案した Translation-Invariant (TI) denoising を利用した TIPSH アルゴリズム [7] を用いることでハールウェーブレットへの依存性を弱める工夫を施す。TIPSH アルゴリズムでは、あるルールに従って shift された観測データに HWT を適用し、level-dependent thresholds を利用した閾値法によって得られた推定結果を unshift することで最終的な推定結果を出力する。本研究では文献 [7] と同じく、シフトルールとして $\hat{y}_i = y_{(i+h) \bmod n}$ を使用し、シフトパラメータ h を一定の値に固定するのではなく、 $h = 1, 2, \dots, n$ のように設定を変えながら NDT-WSE を実施する。そして、 h を変えることによって得られる n 個の推定結果の平均を取ったものを出力結果として算出する。

4 数値例

ここでは文献 [9] と同じソフトウェアフォールトデータ (DS1) を用いて、離散型 NHPP モデルにおけるパラメータ推定法の比較を行う。DS1 はテスト終了時刻 $n = 62$ (日)、総累積フォールト数 $x_{62} = 133$ 個を持つフォールト検出個数データである。各パラメータ推定法に基づいて算出された平均値関数 Λ_i と強度関数 λ_i を用いて、実データとの平均二乗誤差 (Mean Square Error; MSE)

$$\text{MSE}_1 = \frac{\sqrt{\sum_{i=1}^n (\Lambda_i - x_i)^2}}{n}, \quad (20)$$

$$\text{MSE}_2 = \frac{\sqrt{\sum_{i=1}^n (\lambda_i - y_i)^2}}{n} \quad (21)$$

と最大対数尤度 (Maximum Log Likelihood; MLL) を比較し、誤差が小さくなるか、または最大対数尤度が大きくなるパラメータ推定法を最良法として定める。

4.1 適合性評価

適合性評価の結果を表 1 に示す。H(\cdot, \cdot) は NDT-WSE, HTI(\cdot, \cdot) は TIPSH アルゴリズムを使用した NDT-WSE を意味し、h, s, ldt はそれぞれ hard thresholding, soft thresholding, level-dependent thresholds を示している。この表から、TI denoising を取り入れた方法は NDT-WSE のみを使用した場合よりも MSE が小さく、かつ MLL が大きいことを観測することができる。こ

表 1: 適合性評価.

DS1	MSE ₁	MSE ₂	MLL
H(h, ldt)	0.24	0.20	-102.54
H(s, ldt)	0.39	0.26	-120.49
HTI(h, ldt)	0.21	0.19	-98.16
HTI(s, ldt)	0.44	0.26	-119.87

表 2: 代表的な離散型 NHPP モデル.

モデル名	平均値関数 Λ_i ($i = 1, 2, 3, \dots$)
幾何分布 (GE)	$\omega\{1 - (1 - p)^i\}$
負の二項分布 (NB)	$\omega\{\sum_{k=1}^i \frac{(r+k-1)!}{(r-1)!k!} p^r (1-p)^{k-1}\}$
離散ワイブル分布 (DW)	$\omega\{1 - p^{i^r}\}$

(注 : $\Lambda_0 = 0$)

表 3: パラメトリック手法との比較.

DS1	MSE ₁	MSE ₂	MLL
H(h, ldt)	0.24	0.20	-102.54
HTI(h, ldt)	0.21	0.19	-98.16
GE(MLE)	0.68	0.31	-116.75
NB(MLE)	1.05	0.32	-112.65
DW(MLE)	0.69	0.31	-116.69
GE(LSE)	0.59	0.31	-142.73
NB(LSE)	0.92	0.32	-137.90
DW(LSE)	0.59	0.31	-142.63

表 4: データ変換を伴う WSE との比較.

DS1	MSE ₁	MSE ₂	MLL
H(h, ldt)	0.24	0.20	-102.54
HTI(h, ldt)	0.21	0.19	-98.16
HA(h, lht)	1.32	0.11	-72.76
HA(h, ut)	2.57	0.32	-145.39
HF(h, lht)	0.20	0.10	-73.67
HF(h, ut)	0.59	0.31	-142.65

これは、ハールウェーブレットの利用による階段状のノイズが TI denoising によって除去された結果であると考えられる。また、ハード閾値法はソフト閾値法より大きい MLL を与えるため、H(h, ldt) と HTI(h, ldt) が NDT-WSE の中でも最も優れている手法として選出することができる。

4.2 従来手法との比較

次に、提案手法と従来手法の比較を行う。まず文献 [9] と同じように、表 2 に示す代表的な離散型 NHPP モデルを扱い、パラメトリック推定法である MLE 及び LSE による推定を行った。その結果を表 3 に示す。MSE₁, MSE₂, MLL のいずれの評価基準を見ても、NDT-WSE は MLE より優れた推定法であることが結論付けられる。特に、NDT-WSE では尤度を最大化する操作を行っていないにもかかわらず、MLE より大きい対数尤度を与えることが可能である。ほぼ同様の結果は LSE との比較結果にも見受けられる。パラメトリックなモデルに LSE を適用したとして

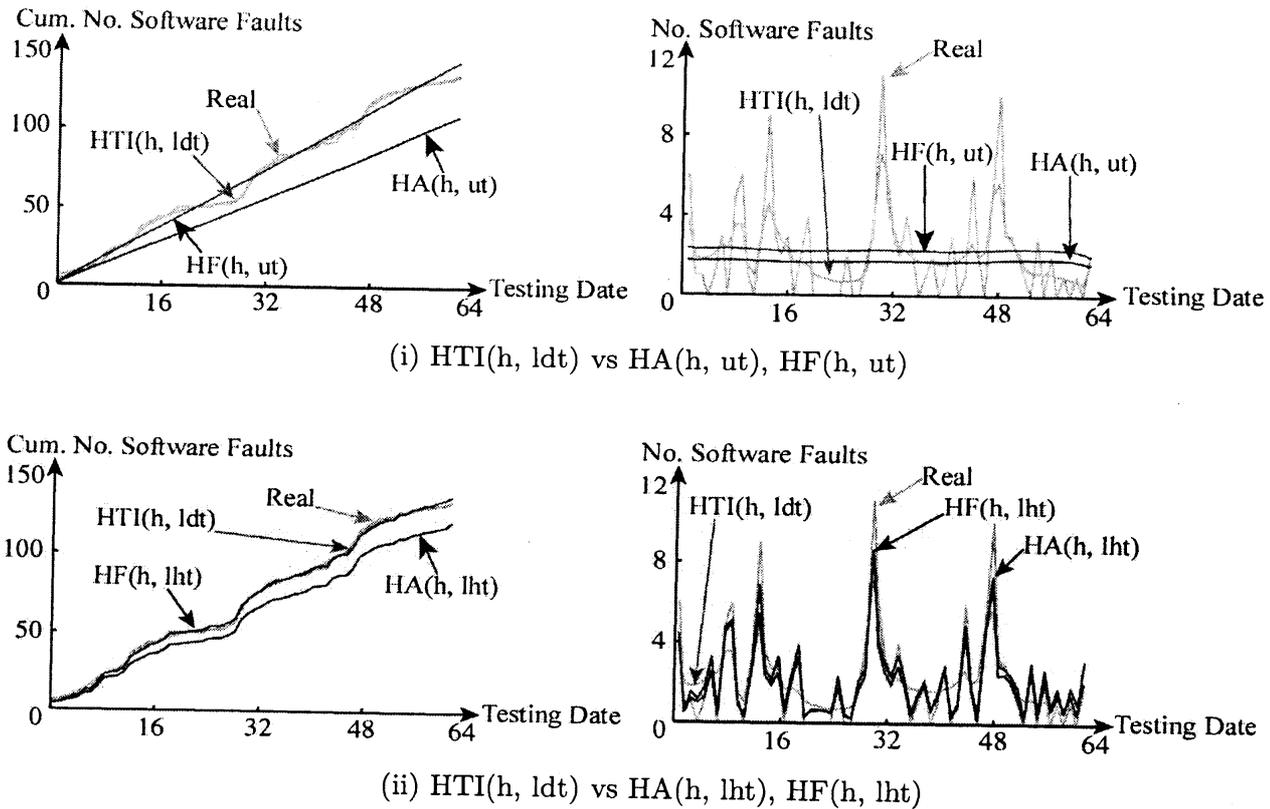


図 1: ウェーブレット縮小推定による推定値の振舞い (DS1).

も、NDT-WSE の推定精度を超えることは非常に稀であり、離散型 NHPP モデルの平均値関数と実データの距離を最小にするという目的に対しても、ウェーブレットに基づいた方法は非常に優れていると結論付けることができる。

次に、Xiao and Dohi [9] が提案した DT-WSE との比較結果を表 4 に示す。HA, HF はそれぞれ Anscombe Transform と Fisz Transform を利用した DT-WSE を表しており、lht, ut はそれぞれ leave-out-half threshold, universal threshold の略である。この表から、 $H(h, ldt)$ と $HTI(h, ldt)$ は $HA(h, ut)$ と $HF(h, ut)$ よりも小さい MSE と大きい MLL を与えることが分かる。TIPSH アルゴリズム及び level-dependent thresholds の利用により階段状のノイズを除去することができ、かつデータ変換が原因で生じる oversmoothing を回避できたと考えられる。しかし、 $HA(h, lht)$ や $HF(h, lht)$ の推定精度には及ばず、シフトルールやシフトパラメータの調整などに関する課題が依然として残されている。

図 1 は、DS1 における推定結果を視覚的に表したものである。図中の (i) は提案手法と $HA(h, ut)$, $HF(h, ut)$ の比較である。数値的な評価と同様に、データ変換を伴う手法では、ソフトウェアフォールトデータの特徴を捉える能力がやや劣っていることが分かる。一方、(ii) に示した累積値の比較では、 $HTI(h, ldt)$ は $HA(h, lht)$ や $HF(h, lht)$ と同様な推定能力を持つように見受けられる。しかしながら、強度に着目し、各テスト日当たりのフォールト検出数の変動を捉えることに主眼を置けば、さらに推定精度を向上させる必要がある。

5 まとめと今後の課題

本研究では、データ変換を伴わないウェーブレット縮小推定を離散型 NHPP モデルに適用し、ウェーブレットに基づいた新しいソフトウェア信頼性モデルの評価法を開発した。実データを用いた検証実験において、パラメトリック推定法である最尤法や最小二乗法に基づいた評価よりも、提案手法は単位テスト時間当たりの検出フォールト数を高い精度で推定できることが示された。また、ノンパラメトリック推定法に関する先行研究と比べても、提案手法の有効性が示された。

今後は、Translation-Invariant denoising におけるシフトルールを選択やシフトパラメータの調整を行うなど、推定精度のさらなる向上が課題として挙げられる。

参考文献

- [1] P. Besbeas, I. De Feis and T. Sapatinas, “A comparative simulation study of wavelet shrinkage estimators for Poisson counts,” *International Statistical Review*, **72** (2), pp. 209–237, 2004.
- [2] R. R. Coifman and D. L. Donoho, “Translation - Invariant Denoising”, *In Wavelets and Statistics*, **103**, pp. 125–150, Springer-Verlag, New York, 1995.
- [3] D. L. Donoho and I. M. Johnstone, “Ideal spatial adaptation by wavelet shrinkage,” *Biometrika*, **81** (3), pp. 425–455 (1994).
- [4] D. L. Donoho, I. M. Johnstone, G. Kerkycharian and D. Ricard, “Wavelet shrinkage: asymptopia? (with discussion),” *Journal of the Royal Statistical Society: Series B*, **57**, pp. 301–337 (1995).
- [5] M. R. Leadbetter, G. Lindgren and H. Rootzén, *Extremes and Related Properties of Random Sequences and Processes*, Springer-Verlag, New York, 1983.
- [6] M. R. Lyu (ed.), *Handbook of Software Reliability Engineering*, McGraw-Hill, New York, 1996.
- [7] E. D. Kolaczyk, “Non-parametric estimation of gamma-ray burst intensities using Haar wavelets,” *The Astrophysical Journal*, **383** (1), pp. 340–349, 1997.
- [8] X. Xiao and T. Dohi, “On equilibrium distribution properties in software reliability modeling,” *Proceedings of 4th International Conference on Availability, Reliability and Security (ARES'09)*, pp. 158–165, IEEE CS Press, 2009.
- [9] X. Xiao and T. Dohi, “Wavelet-based approach for estimating software reliability,” *Proceedings of 20th International Symposium on Software Reliability Engineering (ISSRE'09)*, pp. 11–20, IEEE CS Press, 2009.
- [10] S. Yamada and S. Osaki, “Discrete software reliability growth models,” *Applied Stochastic Models and Data Analysis*, **1** (1), pp. 65–77, 1985.