

# 定量的・定性的メトリクスに基づいた要求定義プロセスにおける 欠陥数推定モデル

野中 誠<sup>†</sup> 阿部 玲子<sup>‡</sup> 矢野 雅嗣<sup>‡</sup>

東洋大学経営学部<sup>†</sup> 三菱電機(株)情報総合技術研究所<sup>‡</sup>

## 1. はじめに

ソフトウェアレビューの十分性を評価する際に、レビュー工数密度や指摘密度などの定量データがしばしば用いられる[1]。しかし、これらのメトリクスだけでは、レビューによって欠陥が十分に除去されていることを高い精度で評価できない。その理由のひとつは、これらのメトリクスの値は欠陥混入数によって影響を受けるが、欠陥混入数という未知の値を推測できていないためである。

欠陥混入数の予測に関する先行研究は、その多くが開発プロセス全体での欠陥混入数を対象としており、要求定義など特定のプロセスを対象としたものは少ない。衣笠らは、ソフトウェア方式設計工程を対象に、当該工程での欠陥混入数を予測する回帰モデルを実データから導出している[2]。このモデルでは、開発規模と仕様変更可能性の2つが説明変数として用いられている。しかし、上流工程での欠陥混入にはこれ以外の要因も影響することが指摘されている[3][4][5]。欠陥混入数を高い精度で予測するには、これらの要因も考慮する必要がある。

一方で、事前情報である欠陥の混入と、事後情報であるレビュー結果などとの間にある因果関係を掘り所にして、事後情報に基づいて欠陥混入数を推定する方法も先行研究において提示されている。Fentonらは、欠陥混入と除去に関わる因果関係をベイジアンネットワークで表現し、残存欠陥数を推定するモデルを提示している[5]。また、阿部らは、レビュープロセスにおける欠陥混入と除去の因果関係をガウシアンネットワークで表したモデルを提示している[6]。

本稿では、阿部らの研究アプローチを踏襲して、要求定義プロセスにおける欠陥混入および欠陥除去について先行研究で提示された知見を取り入れ、阿部らが提示した欠陥数推定モデルを洗練させた結果を報告する。

## 2. 欠陥数推定モデル

### 2.1 ガウシアンネットワーク

ベイジアンネットワークとは、各ノードがそれぞれ確率変数で表され、ノード間の矢印が因果関係を表し、その因果関係が条件付き確率で定量化されている有向非循環グラフのことである。特に、各ノードを正規分布に従った連続型の確率変数で表したものをガウシアンネットワークと呼ぶ。レビュー指摘密度やレビュー工数密度はいずれも連続変数であるため、ガウシアンネットワークを用いるのが適している。また、欠陥混入数は本質的には離散変数であるが、連続変数と見なして扱っても実用的には問題ないと考える。このため、本研究ではガウシアンネットワークを採用する。

### 2.2 ベースとなる欠陥数推定モデル

図1は、阿部らが提示した欠陥数推定モデル[6]である。実線のノードは質的変数、二重線のノードは連続型の量的変数を表している。矢印は因果関係であり、著者らの経験的知識に基づいてモデル化されている。ノードの選定にあたっては、実務での利用性を考慮し、定量的に測定可能な変数および、定性的ではあるが何らかの基準によって測定可能な変数が選ばれている。

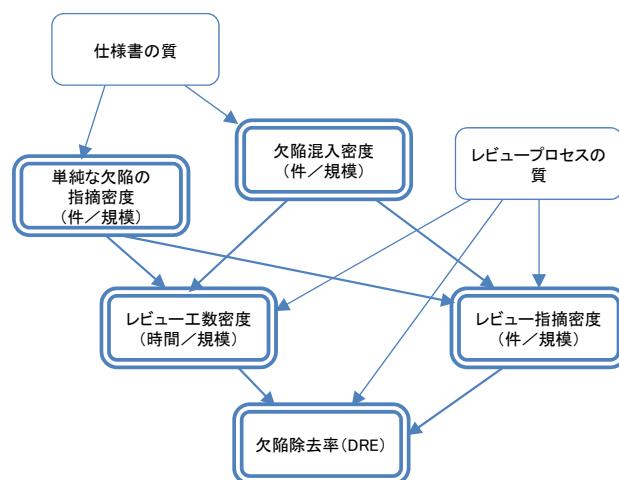


図1 ベースとなる欠陥数推定の因果モデル[6]

A Defect Estimation Model for Requirement Definition Process by using Quantitative and Qualitative Metrics

<sup>†</sup>Makoto Nonaka, Toyo University

<sup>‡</sup>Reiko Abe and Masatsugu Yano, Mitsubishi Electric Corp.

### 2.3 先行研究で提示されている知見

文献[6]では、変数同士の因果関係の定式化にあたって、著者らの経験的な知識に基づいて任意に設定されているため、妥当性の客観性において課題がある。そこで、先行研究において提示されている有用な知見を取り入れ、モデルの洗練を試みる。その知見の一部を次に示す。

- (1) 要求インスペクションでの除去欠陥数は、レビュー工数と正の相関があるが ( $r = 0.72$ )、ロギングミーティング工数単独では相関がなく、事前準備工数とのみ正の相関がある ( $r = 0.70$ )。ただし規模の影響を除外した偏相関係数は  $r = 0.41$  である[3].
- (2) 要求インスペクションでの除去欠陥数は、レビュー対象成果物の規模と正の相関があるが、もっとも有用な規模メトリクスは機能要求数である ( $r = 0.74$ )。ただし事前準備工数の影響を除外した偏相関係数は  $r = 0.51$  である[3].
- (3) 要求インスペクションでの除去欠陥数は、インスペクタの違いだけでなく、成果物の作成者によって異なる[3].
- (4) インスペクションでの除去欠陥数に影響する要因は、インスペクション時間、インスペクタの数、インスペクタの経験、事前準備工数である[4].
- (5) 仕様書の質は、成果物の複雑度、リソースの質、仕様の安定性が影響する[5].

文献[3]では、これらのほかにも機能要求 1 項目あたりの除去欠陥数の平均など、欠陥数推定モデルの初期値を与えるのに有用なデータが提示されている。これらの値もモデルの洗練において利用する。なお、留意すべき事項として、図 1 のモデルのノードでは工数密度や指摘密度などのように規模で正規化した変数を用いているが、上記の知見では正規化されていない変数について言及している点である。

### 3. 洗練後のモデル

前項で示した知見に基づいて、欠陥数予測モデルを洗練した。図 2 にその結果を示す。太字矢印で示した部分が主な改訂箇所である。

知見(1)に基づき、「レビュー工数密度」と「レビュー指摘密度」のあいだに因果関係の矢印を加えた。また、知見(2)に基づき、規模はすべて「機能要求の数」とすることを想定した。

「仕様書の質」を構成する下位概念は、知見(3)および(5)に基づいて構成できると考える。なお、図 1 にあった「単純な欠陥の指摘密度」は、「仕様書の質」を構成する下位概念に含まれると考えられるため、図 2 では削除している。

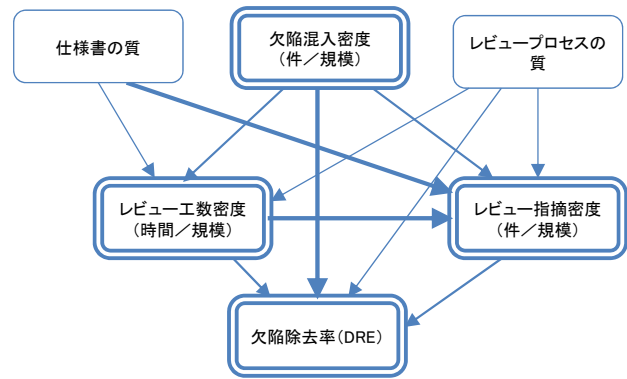


図 2 洗練後の欠陥数推定の因果モデル

それぞれの因果関係における定式化についても、その一部は知見の利用によりモデル化ができる。本稿では紙幅の都合により割愛する。

### 4. おわりに

本稿では、文献[6]が示した欠陥数推定モデルをもとに、先行研究で示された知見を加えることでモデルを洗練させ、要求定義プロセスにおける欠陥数推定モデルを提示した。先行研究の知見を流用することにより、モデルの妥当性を向上させることができたと考える。今後の課題として、実データによるモデルの検証と、モデルを適用する組織のコンテキストに合わせてモデルを適応させる方法の検討が求められる。

### 参考文献

- [1] 野中誠：ソフトウェアインスペクションの効果と効率, 情報処理, vol.50, no.5, pp.385-390, 2009.
- [2] 衣簾宏和, 野中誠, 阿萬裕久：工程別の欠陥埋め込み数の予測—ソフトウェア方式設計を対象として—, ウィンターワークショップ 2012・イン・琵琶湖論文集, p.2, 2012.
- [3] Laitenberger, O., Beil, T. and Schwinn, T., An Industrial Case Study to Examine a Non-Traditional Inspection Implementation for Requirements Specifications, *Empirical Software Engineering*, vol.7, pp.345–374, 2002.
- [4] Suma V. and Gopalakrishnan Nair T.R., Defect Management Strategies in Software Development, *Recent Advances in Technologies*, Strangio, M. A. ed., pp.379-404, In Tech, 2009.
- [5] Fenton N, Krause P, Neil M, "Probabilistic Modelling for Software Quality Control", *Journal of Applied Non-Classical Logics* 12(2),pp.173-188, 2002.
- [6] 阿部玲子, 矢野政嗣, 野中誠：レビュープロセスを対象としたガウシアンネットワークによる欠陥数推定モデル, ウィンターワークショップ 2012・イン・琵琶湖論文集, p.2, 2012.