

論文の要旨

題目 A Study on Refined Neural Network Approach with Data Transformation for Software Fault Prediction

(ソフトウェアフォールト予測のためのデータ変換を用いたニューラルネットワークアプローチに関する考察)

氏名 Momotaz Begum

Software fault prediction is one of the most fundamental techniques in software fault management, and is used to assess the software product reliability and to control the development process. For the software fault prediction, software reliability growth models (SRGMs) have been widely used during the last four decades. But it is worth mentioning that the best SRGM with the highest goodness-of-fit to the past observation does not always possess the best predictive performance to the future (unknown) software fault detection pattern. Many researchers proposed artificial neural networks (ANNs) for the software fault prediction. Under the assumption that the software fault count data follows a Poisson process with an unknown mean value function, we transform the underlying Poisson count data to the Gaussian data via five data transformation methods. In this thesis, we focus on the prediction of the cumulative number of software faults, predictive interval, and software release decision to stop the software testing, and propose a refined neural network approach with the grouped data, where the multi-stage look ahead prediction is carried out with a simple multilayer perceptron neural network with multiple outputs.

In this thesis, we concern the software fault prediction using a multilayer-perceptron neural network, where the underlying software fault count data is transformed to the Gaussian data, by means of the well-known five data transformation methods. More specially, we mainly consider the long-term prediction of the number of software faults, and propose a refined neural network approach with the grouped data, where the multi-stage look-ahead prediction is carried out with a simple multilayer perceptron neural network with multiple outputs. In details, we discuss two different research topics; one-stage look-ahead prediction and multi-stage look-ahead prediction.

In Chapter 2, we focus on a prediction problem with the common multilayer perceptron neural network of the cumulative number of software faults in sequential software testing. We apply the well-known back propagation algorithm for feed forward neural network architectures. We also discuss not only the point estimation but also the sensitivity of the neural network architecture on input neuron and hidden neuron by applying the well-known “rules of thumb” techniques. It is revealed that the rule of thumb is rather accurate to obtain the nearly optimal network architecture for real failure data analyses.

In Chapter 3 and Chapter 4, we pay our attention to the software fault prediction and prediction interval for long term. Here we study the long-term prediction of the number of software faults, and propose a refined neural network approach with the grouped data, where the multi-stage look-ahead prediction is done with a simple multi-layer perceptron neural network with multiple outputs. Under the assumption that the software fault count data follows a Poisson process with an unknown mean value function, we transform the underlying Poisson count data to the Gaussian data via five data transformation methods. Next, we predict the long-term behavior and derive the predictive interval of the cumulative number of software faults in sequential software testing by the refined neural network. On the other hand, nonhomogeneous Poisson process (NHPP)-based SRGMs are widely used for modeling the detection of software faults in software testing. By comparing our proposed models with conventional ones, we show the utility of our neural network models. It is exposed which method is an appropriate one in the both viewpoints of point estimation and interval estimation, throughout our simulation experiments and real failure data analyses.

In Chapter 5, we concern the optimal software release time which minimizes the relevant expected software cost via a refined neural network approach with the grouped data, where the multi-stage look-ahead prediction is performed with a simple three-layer perceptron neural network with multiple outputs. To our best knowledge, there have no research result on the optimal software release problems for long-term prediction via a refined neural network approach. To analyze the software fault count data which follows a Poisson process with unknown mean value function, we transform the underlying Poisson count data to the Gaussian data by means of one of five data transformation methods, and predict the cost-optimal software release time via a neural network. We compare our neural network approach with the common NHPP -based SRGMs. Finally, we conclude the thesis with some remarks in Chapter 6.