PREFACE

The research work presents a systematic approach to identify the temporal patterns that exist in various forms like Bugs, Clones and Failures across different versions of a software application. The thesis used advanced time series analysis for modelling these increasing and decreasing patterns of a particular software characteristic (Bugs, Clones or Failure Intervals). This will give effective feedback to software developers and testing team in advance. This will reduce the effort invested in testing and maintenance. This will help software managers to decide on resource allocation and effort investments. Predicting failure intervals in advance will also improve the software reliability if the developers take timely corrective measures.

The 1st part of this work presents a time series approach to predict the temporal bug numbers and patterns across different versions of a software application. We have applied advanced time series modelling techniques to model the temporal patterns in the bug numbers. Advance knowledge about bug numbers will help the software managers to decide on resource allocation and effort investments. The developers will be aware of the number of bugs in advance and can take effective steps to reduce the number of bugs in the new version. The end user can decide on adopting a particular software application among a variety of applications by knowing the bug growth patterns of the particular software application. The experiments are validated against bug number data collected from bug repositories of Debian, Eclipse and Mozilla.

The 2nd part of this work presents a systematic approach for modelling the evolving clones across different versions of an open source software application. It is observed that MOGA-NN based approach is most promising and effective in predicting clone evolution. Clone detection is useful for reducing the Corrective Maintenance and Preventive Maintenance which involves modification of code content to solve and prevent problems in the

software respectively. Because if we can identify and detect the cloned areas, the defect in all the similar code fragments can be resolved at once. The software clone evolution prediction is immensely helpful in Perfective Maintenance and Adaptive Maintenance because the effort required to evolve a software is also dependent on the amounts of cloned contents in the software. The experiments are validated against clone information data collected from source code of ArgoUML.

The last part of this research presents an application of time-series modelling for reliability prediction based on TBF pattern. From the Results, it is observed that Hybrid ARIMA (ARIMA +ANN) as a good predictor of software reliability. An interval-based approach using MOGA–NN and ELM eliminated the problem of associated errors due to uncertainties in the model parameters and noise in the input data. Accurate prediction of software reliability will decrease the risk of failures. The experiments are validated against failure interval data collected from CSIAC software reliability dataset.

In summary, time series approach is used in the thesis to improve Software Bug Prediction, Clone Evolution Prediction, and Software Reliability Prediction. Advanced Time Series Modeling using Statistical and Machine Learning Approaches have been used for modelling the temporal patterns. The dataset used are obtained from Open Source Software Repository and also Closed Source Software Repository. The results are evaluated using Standardized Evaluation Techniques and compared against other reported methods.