A Software Reliability Model Based on a Geometric Sequence of Failure Rates

Stefan Wagner
Software & Systems Engineering
Technische Universität München
Germany
wagnerst@in.tum.de
(joint work with Helmut Fischer, Siemens AG)

June 7, 2006

Introduction

Background and Motivation

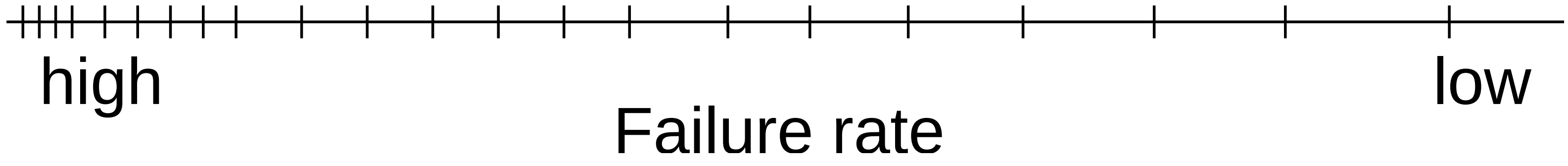
- *Software reliability engineering* is an established area of software engineering concerned with the measurement and improvement of reliability
- For the measurement typically stochastic models are in use
- They model the failure process and use other software metrics or failure data for the parameter estimation
- They are able to
  - estimate the current reliability and
  - predict future failure behaviour
- In Milller (1986) it is suggested that geometric rates (a geometric relationship between the failure rates) are possible
- This behaviour was observed by NASA as documented in Nagel et al. (1982, 1984)

Problem

- The underlying problem is the classical *when to stop testing* problem of Siemens products
- For software reliability engineering in general, there is still a need for accurate models for different environments and projects
- Specifically, in the analysed area at Siemens, a geometric sequence of the failure rates of faults was observed
- How can we model that and use it to analyse the test and production failure behaviour?

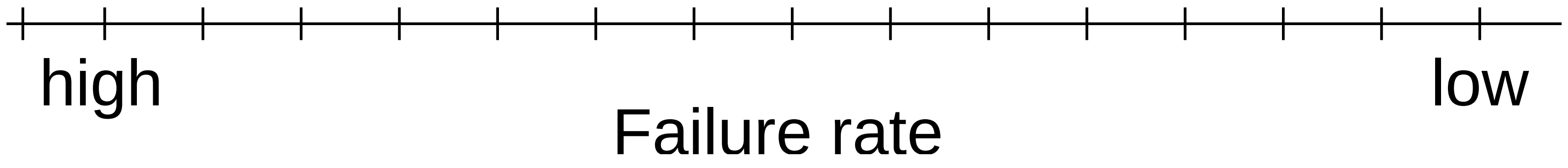
The Model

- Failure rate of a fault: The probability that an existing fault will result in an erroneous behaviour of the system during a defined time slot or while executing an average operation
- Assumption: There are many more faults with high failure rates than with low failure rates
- Hence the distribution of failure rates of faults looks like this:



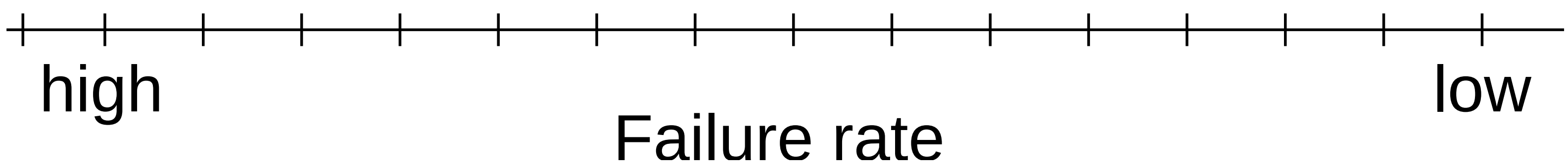
Basics

- Failure rate of a fault: The probability that an existing fault will result in an erroneous behaviour of the system during a defined time slot or while executing an average operation
- Assumption: There are many more faults with high failure rates than with low failure rates
- Using a logarithmic scale:



Basics

- Failure rate of a fault: The probability that an existing fault will result in an erroneous behaviour of the system during a defined time slot or while executing an average operation
- Assumption: There are many more faults with high failure rates than with low failure rates



Hence, the failure rates can be approximated as follows:

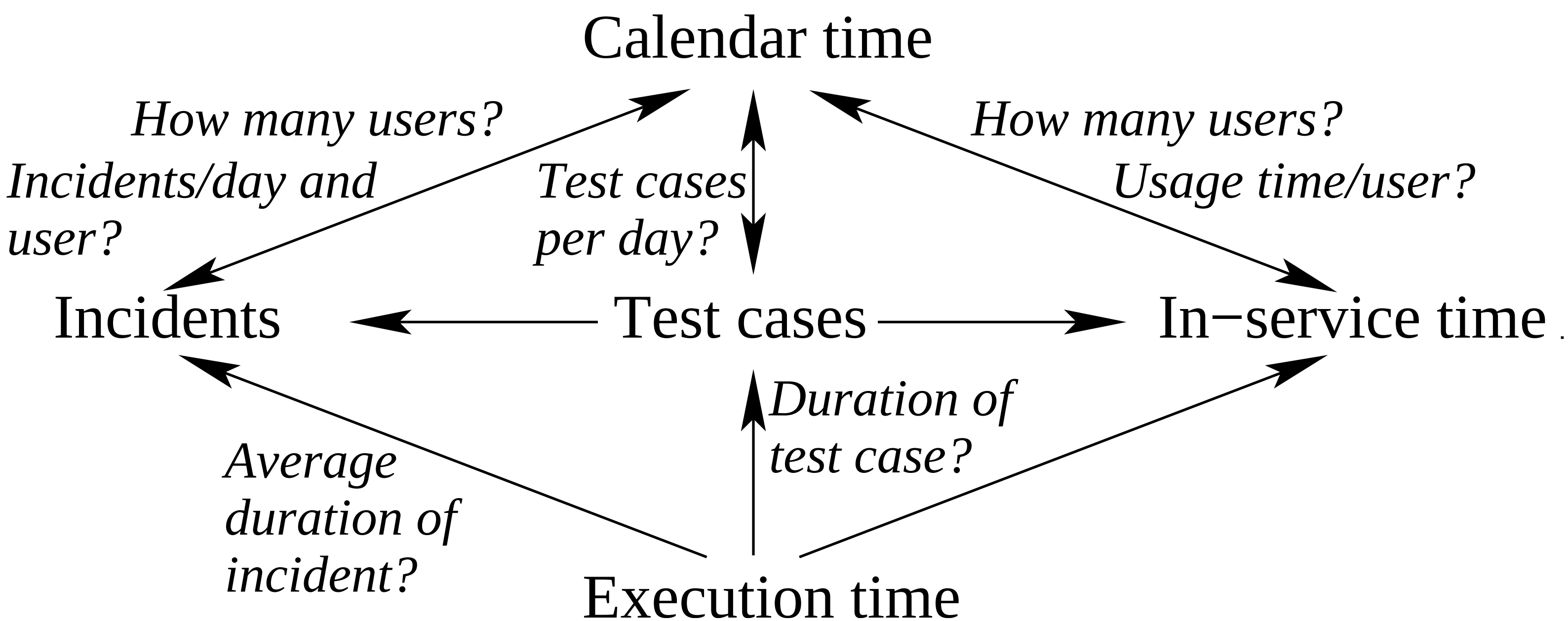
$$p_n = p_1 \cdot d^{(n-1)}, \tag{1}$$

where p_n is the failure rate of the n-th fault, p_1 the failure rate of the first fault, and d is a project-specific parameter.

Quantities

- We assume a geometrical distribution of the failure occurrences
- Based on this, we can derive all usual quantities
  - Mean number of failures (μ)
  - Failure intensity (λ)
  - Time to reach a certain reliability level (Δt)
- Based on *incidents* as time
- An incident is a typical usage task of the system

Time Component



Evaluation

Quality of Assumptions

- Main Assumption: geometrical relationship of failure rates of faults

Quality of Assumptions

- Main Assumption: geometrical relationship of failure rates of faults
  supported by Nagel et al. (1982, 1984), tendency also by Adams (1984)

Quality of Assumptions

- Main Assumption: geometrical relationship of failure rates of faults
  supported by Nagel et al. (1982, 1984), tendency also by Adams (1984)
- The occurrence of a failure is geometrically distributed

Quality of Assumptions

- Main Assumption: geometrical relationship of failure rates of faults
  supported by Nagel et al. (1982, 1984), tendency also by Adams (1984)
- The occurrence of a failure is geometrically distributed
  fits to the considered discrete incidents

Quality of Assumptions

- Main Assumption: geometrical relationship of failure rates of faults
  supported by Nagel et al. (1982, 1984), tendency also by Adams (1984)
- The occurrence of a failure is geometrically distributed
  fits to the considered discrete incidents
- Infinite number of faults

Quality of Assumptions

- Main Assumption: geometrical relationship of failure rates of faults
  supported by Nagel et al. (1982, 1984), tendency also by Adams (1984)
- The occurrence of a failure is geometrically distributed
  fits to the considered discrete incidents
- Infinite number of faults
  Imperfect debugging

Applicability

- Real application only in telecommunications domain
- During evaluation of predictive validity also used on data from other domains
- In principle usable during system test, field trial, and field use

- Concept of *incidents* does not fit to all domains

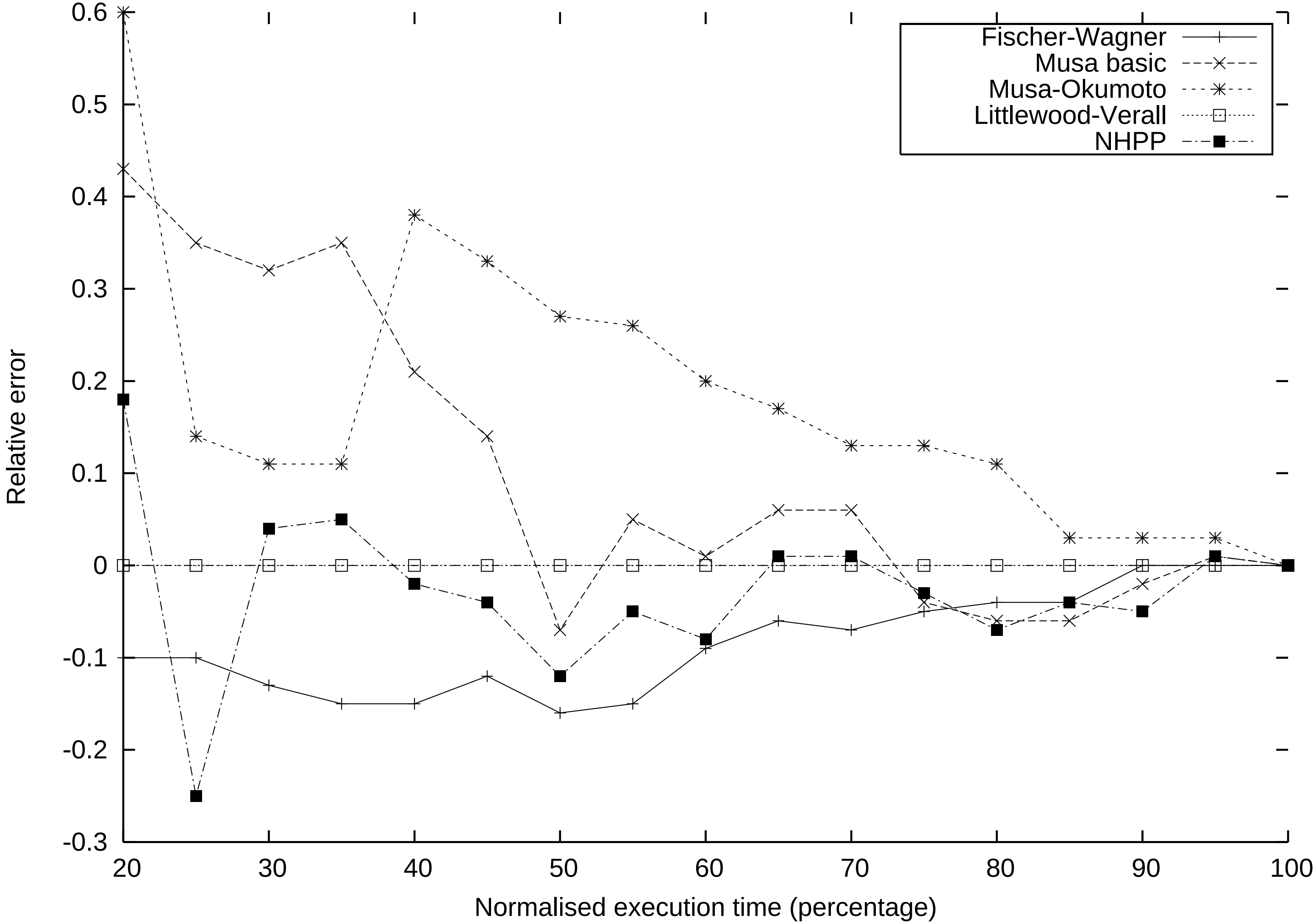
Simplicity

- Consideration of faults and failures complicated
- Each fault has a different failure distribution
- Infinite number of faults difficult for tool support
- The two parameters have an interpretation:
  - p_1: highest failure probability
  - d: system complexity

Predictive Validity – Approach

- Comparison of predicted and measured numbers of failures and several points in time
- Relative error: $(\hat{\mu}(t_q) - q)/q$
- Models:
  - Musa basic
  - Musa-Okumoto
  - Littlewood-Verall
  - NHPP
- Data Sets:
  - Telecommunication systems from Siemens
  - DoD DACS data

Predictive Validity – Results



Conclusions

Summary

- Software reliability model based on a geometric series of the failure rates of the faults
- Suggested by theory and practice
- Predictive validity similar to existing models
- More consistent over different projects

Future Work

- Early determination of model parameters by correlation with other measures
- Improved time component
- Further evaluation on other projects
- Improved tool support