Seminar 12 - Software Reliability Growth Models

PV260 Software Quality

Stanislav Chren, Václav Hála

12. 5. 2015

Software Reliability

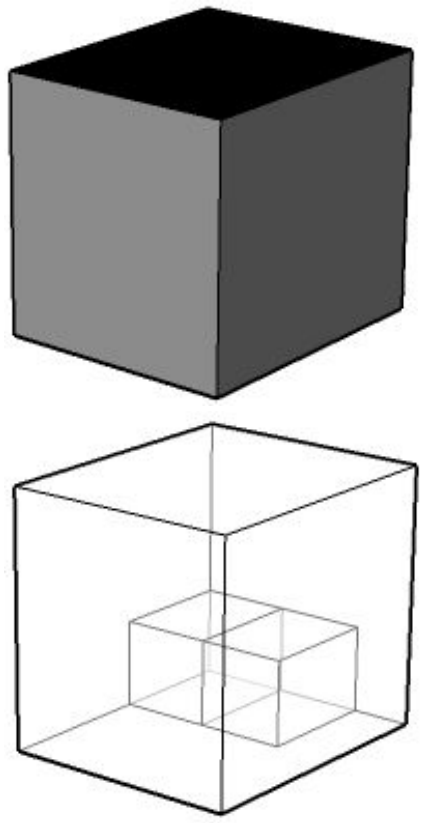
- **IEEE 1633-2008:**
  Probability of failure-free operation for a specified period of time in a specified environment
- **In terminating systems:**
  Probability of failure-free operation per demand/request
- **In continuously running systems:**
  Failure intensity during specified time interval

Software Reliability Parameters

- **Probability of failure (reliability)**
- **Constant failure rate** λ:
  Number of failure occurrences per unit of time.
- **Time-dependent failure intensity** $\lambda(t)$:
  Rate of change of expected number of failure with respect to time.

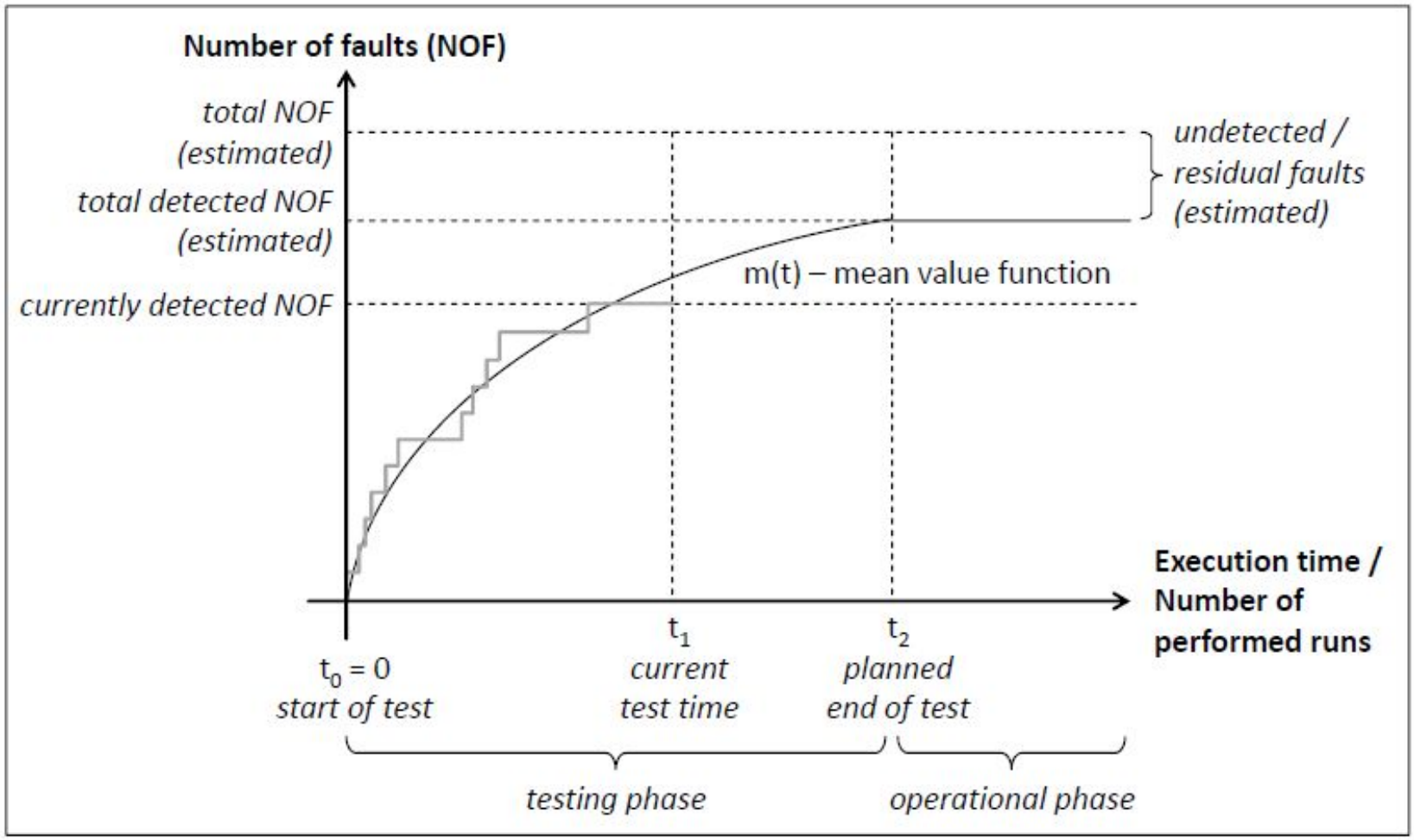
Software Reliability Models

- **Black-box**
  - Software Reliability Growth Models
  - Black-box testing
  - ...
- **White-box**
  - State-based
  - Path-based
  - Additive
  - Fault-trees



Software Reliability Growth Models

- *Models based on estimation of the number of failures/faults and failure intensity based on the historical data*
- $m(t) = a(1 - e^{-bt})$, $\lambda(t) = \frac{dm(t)}{dt}$



SRGM Assumptions

- Faults are repaired immediately when they are discovered
- Fault repairs are perfect
- No new code is introduced during testing
- Defects are only reported by the product testing group
- Each unit of time is equivalent
- Tests represent operational profile
- Failures are independent

SRGM Analysis Steps

1. **Data selection**
   Type, sources, time units
2. **Model selection**
   Laplace's trend test, ...
3. **Parameter estimation**
   Maximum likelihood, Minimum least square,...
4. **Goodness-of-fit**
   R^2 test, Kolmogorov-Smirnov test, Chi-square test
5. **Failure data prediction**

SRGMs and Issue Tracking

In case the dedicated testing data about failure occurrence are not available, they can be sometimes substituted with the bug reports. In order to mitigate the violation of the model assumptions, the following filtration criteria are often used:

- Specific version of the software
- Specific components
- Without duplicate reports
- Fixed/resolved status
- High severity

Tools

- **CASRE**
  www.openchannelsoftware.com/projects/CASRE_3.0
- **SMERFS**
  www.slingcode.com/smerfs/
- **RGA**
  www.reliasoft.com/rga/index.htm

Task - Mozilla SRGM Analysis

1. Open the Bugzilla (`https://bugzilla.mozilla.org/`) and select the advanced search
2. Select the `Core` product and all its components
3. Browse the bug reports based on the following criteria:
   - belong to a specific version (e.g. 30)
   - at least MAJOR severity
   - RESOLVED/CLOSED status
   - FIXED resolution
4. If you obtain less than 50 entries, try different version or relax the severity option (use severities normal and above)
5. Order the reports according to the date and note down the amount of bug reports per selected time periods (e.g. two-week, month or two-months))

Task - Mozilla SRGM Analysis

6. Start the Matlab and create the time and fault count vectors.
7. Construct the vector of cumulative number of fault counts. You can use the `cumsum(X)` function.
8. Create plots for the fault count data and cumulative fault count data. In Matlab you can use the `plot(X,Y,'-ro')` function.
9. Determine the trend of the fault counts. In order to use the SRGM, the trend should be negative. Do it both visually and using the Laplace's tred test:
   - $L = \frac{\frac{\sum t_i}{n} - \frac{T}{2}}{T\sqrt{\frac{1}{12n}}}$,

     t_i is the number of time units from the start to the occurrence of the i-th fault

     n is the total number of faults

     T is the total testing time

Task - Mozilla SRGM Analysis

10. If the trend is negative, select 3-5 SRGMs (from the sources in the study materials)
11. Use the Curve fitting application in Matlab to fit the models with the data
    - The X data is the time vector
    - The Y data is the cumulative fault count vector
    - Select the custom equation option and input the mean value function of the selected SRGM
    - In Fit options you can adjust starting values of the parameters and their boundaries
12. Determine the best model based on the goodness-of-fit statistics (e.g. the R^2 statistics)
13. (Optional) Plot the multiple SRGMs in the same plot with the input data

Task - Mozilla SRGM Analysis

14. Calculate the time-dependent failure intensity function of the best SRGM as the derivation of the mean value function
    - In Matlab:

      ```
      syms a,b,t...
      f = ...
      diff(f,t)
      ```

15. Use the resulting mean value function and the failure intensity function to predict the future number of faults and failure intensity.