Confidentiality and Real Errors: A Contradiction?
Dieter Hein and Manfred Reitenspiess
Fujitsu Siemens Computers, Munich

Summary

During industrial software development and deployment, a wealth of data is accumulated, which could be used for the evolution and refinement of methods and tools for error analysis, statistical evaluation of errors, dynamic handling of errors, and the prediction of faults and failures. Unfortunately, this data is always classified as highly sensitive as it contains customer related information, quality and quality assurance related information and gives insights into internal development processes. There is a need for neutralization techniques to overcome these hurdles.

Errors during Development

High quality software, in particular software for highly-reliable, high-availability applications such as communications systems, automotive, Web requires the use of clearly defined processes and strict adherence to the applicable rules.

At Fujitsu Siemens Computers, the quality process for software development is defined in the “method handbook”. The handbook is structured according to the waterfall model, separating the phases requirements engineering, design, development, testing, piloting, and deployment. Progressing software to the next stage, requires the fulfillment of a number of quality criteria, which are jointly checked by the project management team and the quality management team. One of the criteria is the number of errors found (and corrected). Based on a “projection model”, for each milestone the number of expected (to be found) errors is identified in advance. Widely accepted standards (e.g. Boehm), but more importantly, internal error statistics from similar projects are used for defining the error thresholds.

The waterfall model is often not applicable for a number of reasons, an important reason being the quest for optimal quality. For example, human interface design requires the early involvement of users, even if a number of functions are not yet implemented. Also in high availability environments, the early access to a rudimentary system is important to validate theoretical system models and therefore get early reality checks. Therefore, the spiral model has become widely accepted, taking advantage of test and try approaches (using cheap system resources) and the automation of tests. Tools are available which allow the specification of test cases and the evaluation of complete application scenarios. Last but not least, the train model allows more flexibility in the management of release features. Only features which fulfill given acceptance criteria make it into a release, while others may come late or may even be postponed to a later release.

Error projection methods as described above are also applicable to this more dynamic approach to software development. The experience of thousands of man years of software development are maintained and further evolved by the Fujitsu Siemens Computers engineering organization. After finishing the programming phase, all errors are entered, tracked and maintained in a central error tracking tool. Development and QA organization are both responsible for the data entries and changes including the classification of errors (critical, major, minor), their resolution etc. Errors found during programming phases and earlier are documented. The documentation is an integral part of the acceptance criteria to enter the QA phase.

Access to the error database is highly restricted and classified as confidential. Standard analysis tools are available to track handling times or frequent occurrence of errors. However, the evaluation of errors, effects analysis, detailed defects analysis or repeatable behavior require manual efforts.

Errors in Production Mode

As soon as the software is transferred to a production environment (officially released), an additional dimension of complexity is added as costs for error
identification, correction and field repair grow by factors. Some of the main reasons can be:
- Very complex reproduction of failures and cause analysis in lab environments
- Involvement of maintenance personnel requiring additional process steps for repair
- Customer involvement

Fujitsu Siemens Computers applies the same error tracking tool to manage production mode errors as for development errors. The correction of reported production mode errors, however, involves additional process steps, as the development and quality processes have to interface to the maintenance processes.

The integrated view on reported errors and their management is fed back into the error forecast modeling for systems under development, but can also be used to model error expectations in operational systems and apply preventive maintenance. Field error reports are closely watched as their handling can have critical impacts on customers’ business processes. The correlation of field errors, detecting cross effects of errors or increased frequencies of specific errors/error types requires intimacy with the system under evaluation and is often done in semiautomatic fashion.

The attached graphic is shown to clarify the complex relationships between a software development organization and the customers using the software. A number of organizations need to have access to the error database. In case of software partners, additional filters are necessary to protect company proprietary data from access through other entities (partially due to legal requirements, anti-trust law).

**Industrial Academic Cooperation**

Using real data in the development and analysis of system models can considerably increase their usefulness. Additionally, tools for deriving useful statistics from available error data can be improved or newly developed.

A major hurdle in making real data accessible for scientific research is the control over its use. Error data typically contains references to internal details of the related software which needs to be protected from external access. Field error reports always contain references to customer data which needs to be protected due to legal bindings. Overall, external access to the corporate error data base is prohibited as this data is perceived as highly confidential and core to the company.

Use of real error data for academic research will only be possible, if the confidentiality constraints can be overcome.

**Literature**