

Analyse und Konzept zur Verbesserung der statischen Fehlereingrenzung

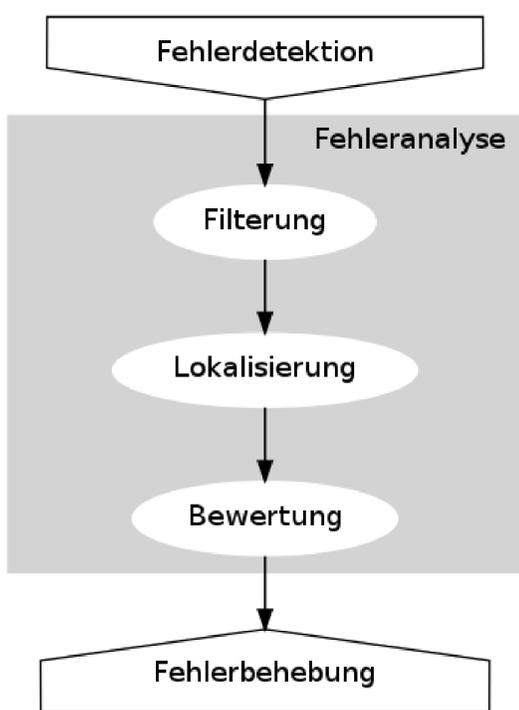
Masterarbeit, vorgelegt von Lars Gohlke

Diese Masterarbeit entstand im Rahmen meiner Tätigkeit bei der Sprd.net AG als Qualitätsingenieur in der Softwareentwicklung. Die Sprd.net AG ist im Bereich des Onlineversandhandels mit u.a. individuell gestaltbaren Textilien tätig.

Aufgabenstellung:

Diese Arbeit beschäftigt sich mit der Erarbeitung der Analyse und eines Konzeptes zur Verbesserung der statischen Fehlereingrenzung mit Hilfe von Softwaremetriken. Die statische Analyse der Tests soll den Fehleranalyseprozess in Bezug auf vorhandene Tests beschleunigen. Auf diese Weise wird die Werkzeugkette des Softwareentwicklungsprozesses wirksam unterstützt und effizienter gestaltet.

- Wie kann der Fehleranalyseprozess nachhaltig und wirksam verbessert werden?
- Wie lassen sich Softwaremetriken in den Fehleranalyseprozess integrieren?
- Kann eine Verbesserung der Effizienz bei der Fehlereingrenzung beobachtet werden?



Allgemeine Darstellung des Prozesses der Fehlersuche

Allgemeine Strategien:

Das Fehlen einer systematischen Fehlereingrenzung macht es notwendig, die Analyse der ausgewählten Tests auf der Grundlage verschiedener Strategien durchzuführen.

1. Sequentielle Strategie

Nach dieser Strategie werden die Fehler aufgrund ihrer Auftretensreihenfolge in der Liste untersucht.

2. Erfahrungsbasierte Strategie

Mit dem Wissen der internen Abläufe werden die wahrscheinlich verantwortlichen Tests ausgewählt.

Metrik "Typdistanz":

Der Kern der Verbesserung wird durch die Einführung einer neuen statischen objektorientierten Strukturmetrik der *Typdistanz* erzielt. Damit erfolgt die Vorsortierung der fehlgeschlagenen Tests bei der verbesserten Strategie mit Hilfe einer Testklassifizierung.

Diese Metrik beschreibt die Entfernung zwischen zwei Klassen (Typen). Diese Entfernung wird hierbei auf einer Absolutskala angegeben.

$$d(V_0, V_1) = \begin{cases} 0 & \text{wenn } V_0 \text{ oder } V_1 \text{ Vertreter eines Typs des Standard-} \\ & \text{sprachumfangs}^{12} \text{ sind} \\ 1 & \text{wenn } V_0 \text{ und } V_1 \text{ zwei Nachbarblätter sind} \\ x & \text{Anzahl der Kanten zwischen } V_0 \text{ und } V_1 + 1 \end{cases}$$

V_0	V_1	$M_{distance}$
de.commons.A	de.commons.B	1
de.commons.A	de.utils.Helper	5 (4+1)
de.commons.A	java.lang.String	0

Einzeltypdistanz für ausgewählte Typen

Die Gesamttypdistanz beinhaltet die Zusammenfassung aller Einzeltypdistanzen eines Tests. Die aktuelle Berechnungsvorschrift der Gesamttypdistanz eines Tests berücksichtigt drei Faktoren:

1. Maximale Einzeltypdistanz

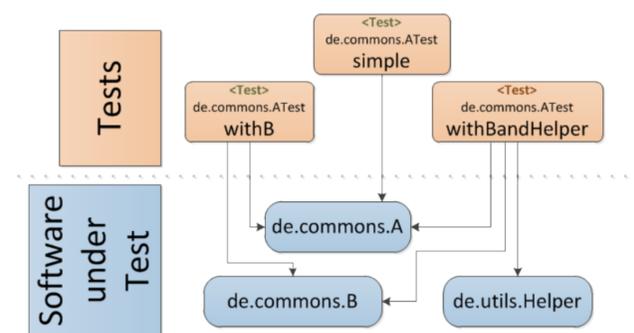
Aufgrund des zentralen Kriteriums der Typdistanz wird das Maximum aller vorhandenen Einzeltypdistanzen am höchsten gewertet.

2. Median der Einzeltypdistanzen

Für eine robuste Ermittlung der Gesamttypdistanz wird der Median berechnet.

3. Anzahl der Anweisungen

Sind die unter 1.) und 2.) genannten Wichtungsfaktoren identisch, ist die einfache Wichtung nach Umfang der Anweisungen ausreichend.



Schematische Darstellung der Einzeltypdistanzen am Beispiel unterschiedlicher Tests

Fehlerfortpflanzung:

Allgemein ist bekannt, dass ein Fehler weitere nach sich ziehen kann. Bei der Softwareentwicklung kann entsprechend dem Prinzip der Fehlerfortpflanzung ein fehlgeschlagener Test u.U. nur als Symptom gewertet werden. Die Fehlerfortpflanzung erzeugt durch die Kopplung der einzelnen Komponenten ein gewisses Fehler-Rauschen in Form weiterer fehlgeschlagener Tests. Es konnte beobachtet werden, dass zusätzliche Fehler, die durch die Fehlerfortpflanzung auftreten, eine größere Gesamttypdistanz aufweisen.

Kleinster Test:

Der kleinste Test beschreibt den Test, der im Rahmen aller fehlgeschlagenen Tests die geringste Gesamttypdistanz aufweist.

Kleinster Fehler:

Ausgehend von den Definitionen für Fehler und kleinster Test können diese, zusammengenommen, auch als kleinster Fehler bezeichnet werden. Sollte in einem kleinsten Test ein Fehler gefunden werden, so wird dieser als kleinster Fehler bezeichnet.

Weiterentwickelte Strategie (bottom-up):

Die Idee dieser Strategie besteht darin, zunächst den kleinsten Test in der Abhängigkeitsstruktur der Tests zu untersuchen. Mit der Annahme, dass der kleinste Test der Fehlerursache am nächsten ist, entfällt ein zusätzliches Überprüfen der größeren Tests.

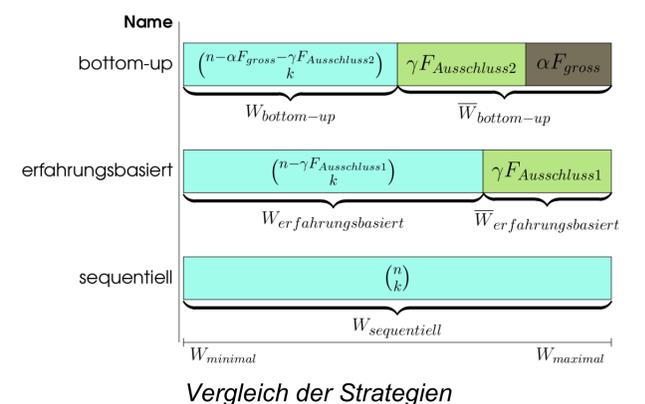
1. Sortierung der Fehler

Die Sortierung der Fehler verfolgt das Ziel, anhand geeigneter Kriterien mit dem *kleinsten Fehler* die Fehlereingrenzung zu beginnen, um auf diese Weise den Aufwand zu minimieren. Diese Sortierung erfolgt idealerweise maschinell am Ende des Testlaufs und ist so optimiert, dass der Aufwand vernachlässigt werden darf.

2. Analyse der Fehler

Die Analyse der kleinsten Tests macht weiterhin manuelle Arbeitsschritte notwendig und bedarf der eingehenden Expertise eines Entwicklers.

- n = Anzahl fehlgeschlagener Tests
- k = Anzahl kleinster Fehler
- $\binom{n}{k}$ = Wahrscheinlichkeit, einen von k kleinsten Fehler aus n fehlgeschlagenen Tests herauszugreifen
- γ = Wichtung der Erfahrung $\{\gamma \in \mathbb{R} \mid 0 \leq \gamma \leq 1\}$
- $F_{Ausschluss}$ = Anzahl der aufgrund von Erfahrung ausgeschlossenen Tests $(0 \leq F_{Ausschluss})$
- \bar{W} = eingesparter Aufwand
- α = Wichtung der Qualität der Filterung $\{\alpha \in \mathbb{R} \mid 0 \leq \alpha \leq 1\}$
- F_{gross} = Anzahl herausgefilterter Tests $(0 \leq F_{gross})$



Fazit:

Thema der vorliegenden Arbeit war die Analyse und Erarbeitung eines Konzeptes zur Verbesserung der statischen Fehlereingrenzung. Dabei ging es darum, einen *leichtgewichtigen heuristischen* Gegenentwurf zu aufwendigen statistischen Verfahren (z.B. Delta Debugging) zu entwickeln. Der Gegenentwurf beinhaltet die Modellierung einer Metrik zur Bestimmung der signifikantesten Tests.

Das Ziel dieser Arbeit bestand darin, die Risiken bei der Weiterentwicklung der Methodik zur Fehlereingrenzung zu reduzieren, den Prozess der Fehlerbehandlung und damit den Prozessschritt der Filterung aufwandsstabiler zu gestalten. Die Instabilität bei der Filterung ergibt sich aus dem Fehler-Rauschen in Folge der Fehlerfortpflanzung und erhöht so den Aufwand zur Lokalisierung der signifikantesten Tests.

Zur Überprüfung bzw. Bestätigung des Konzeptes zur Verbesserung der statischen Fehlereingrenzung wurde an einem Fallbeispiel eine Untersuchung durchgeführt. Aufgrund verschiedener Einflüsse bei der Untersuchung konnten nur erste Erfahrungen für eine weitergehende Studie gesammelt werden. Ein abschließendes Urteil über die Wirksamkeit der verbesserten Methodik ist somit anhand dieser Untersuchung nicht möglich.