IOWA STATE UNIVERSITY

Advances in Time Estimation

Lotfi ben Othmane

- Context: I want to get an accurate budget to fix bugs

- Question: How to get an accurate time estimation to fix the bugs

Question: How to estimate the time to fix bugs?

https://www.google.com/search?source=hp&ei=r7bvXeuLNYXV9AOzuLOgCw&q=estimate+project+time&oq

Google

estimate project time

All Images Videos News Shopping More Settings Tools

About 299,000,000 results (0.90 seconds)

Project Timelines | Made Easy with Smartsheet®

Ad www.smartsheet.com/

The Easiest & Most Complete **Project** Management & **Timeline** Software. Try It Free! 100% Cloud-Based. Make Collaboration Work. Enterprise Ready. Features: Easy, Flexible, Collaborative.

Smartsheet in the News · Make Gantt Charts Online · Take Our Product Tour · Contact Us

Project Management: Time Estimates and Planning

- Step 1: Understand the **Project** Outcome. First you need to fully understand what it is you need to achieve. ...
- Step 2: **Estimate Time**. When you have a detailed list of all the tasks that you must achieve to complete the **project** then you can begin to **estimate** how long each will take. ...
- Step 3: Plan for it Going Wrong.

Project Management: Time Estimates and Planning

https://www.projectsmart.co.uk › project-management-time-estimates-and-pl...

About Featured Snippets Feedback

Estimating Time Accurately - Project Management Skills from ...

https://www.mindtools.com › Project Management › Scheduling

COCOMO MODEL

$$E = a(k\_loc)^b$$

	a	b
Organic	2.4	1.05
Semi Detached	3.0	1.12
Embedded	3.6	1.20

Main limitation of COCOMO:

Unknown size of expected software

What are the exiting methods?

1. Experts judgement
2. Analogous estimating
3. Parametric estimating
4. Three-point estimating
5. Group decision-making estimating

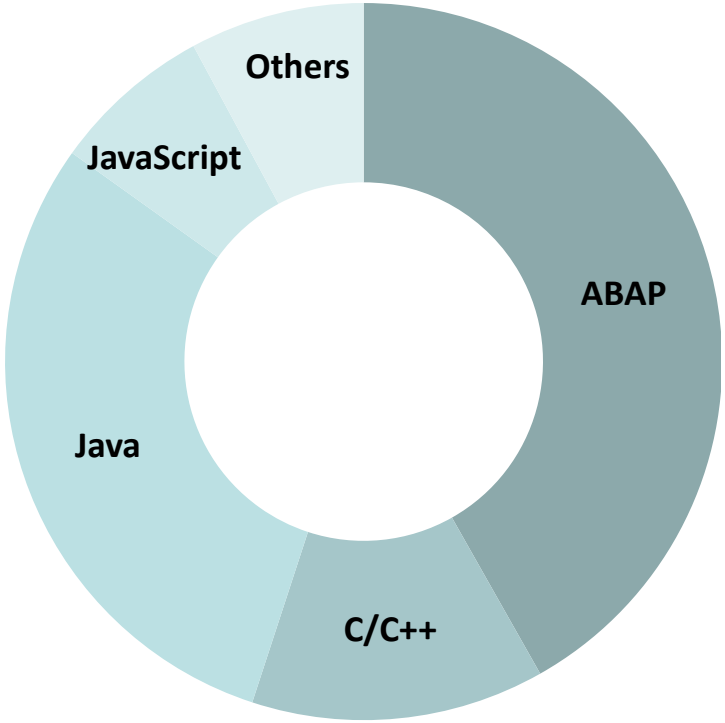
Scientific Research

1. **Research question**
   - Refine to a set of small questions that we can verify

2. **Testing approach**
   - Algorithm implementation
   - Empirical study
   - Simulation

3. **Evidence and argumentation**

Security Testing at SAP



Language	Tool	Vendor
ABAP	CVA (SLIN_SEC)	SAP
C/C++	Coverity	Synopsys
JavaScript, Ruby	Checkmarx	Checkmarx
Others	Fortify	HP

- Mandatory for all products
- Multiple billons lines analyzed
- Constant improvements:
  - tool configuration
  - new tools and methods

Is there a More Accurate Estimation Method?

- Goal: Identify the factors that impact the fixing time
- Method: Interview participants in the vulnerability fixing process
- Result: The major factors that impact the fixing time

Factors of the Vulnerability Fix Time

Factor categories	# of factors	Freq.
Vulnerabilities characteristics	6	9
Software structure	19	10
Technology diversification	3	5
Communication and collaboration	7	8
Availability and quality of information and documentation	9	9
Experience and knowledge	12	11
Code analysis tool	4	4
Other	4	4

Is there a More Accurate Estimation Method?

Challenge: Predict the cost of fixing a given security vulnerability

$\Rightarrow$ Predict the duration of fixing security vulnerabilities?

Let vul_fix_time = $f(x_1, \ldots\ldots x_j)$

Given collected attributes x_i?

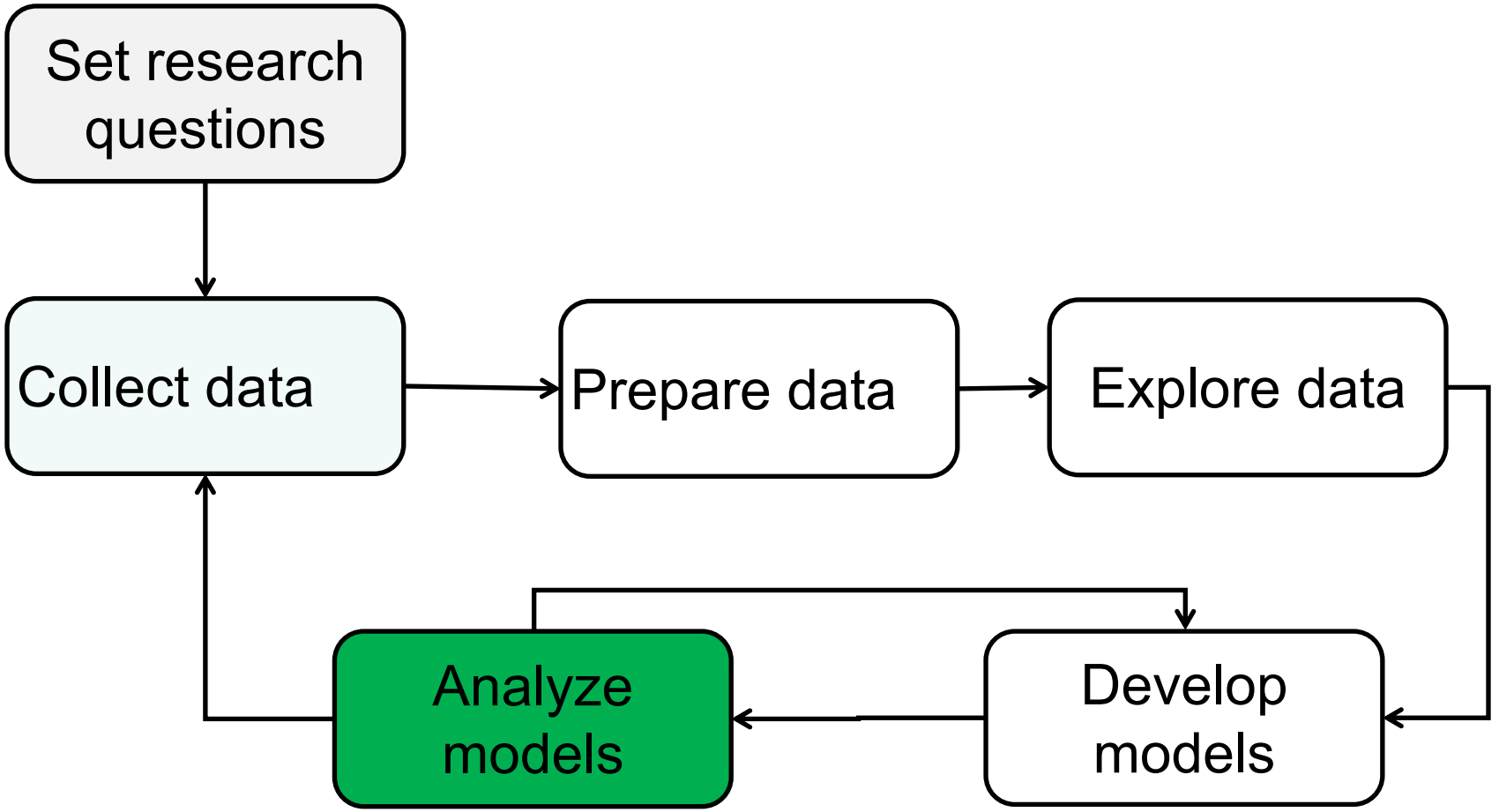
Find out *f*

Data-Sets

1. Static analysis data of ABAP code (Data-set 1)
2. Static analysis data of Java and C Code (Data-set 2)
3. Security messages (Data-set 3)

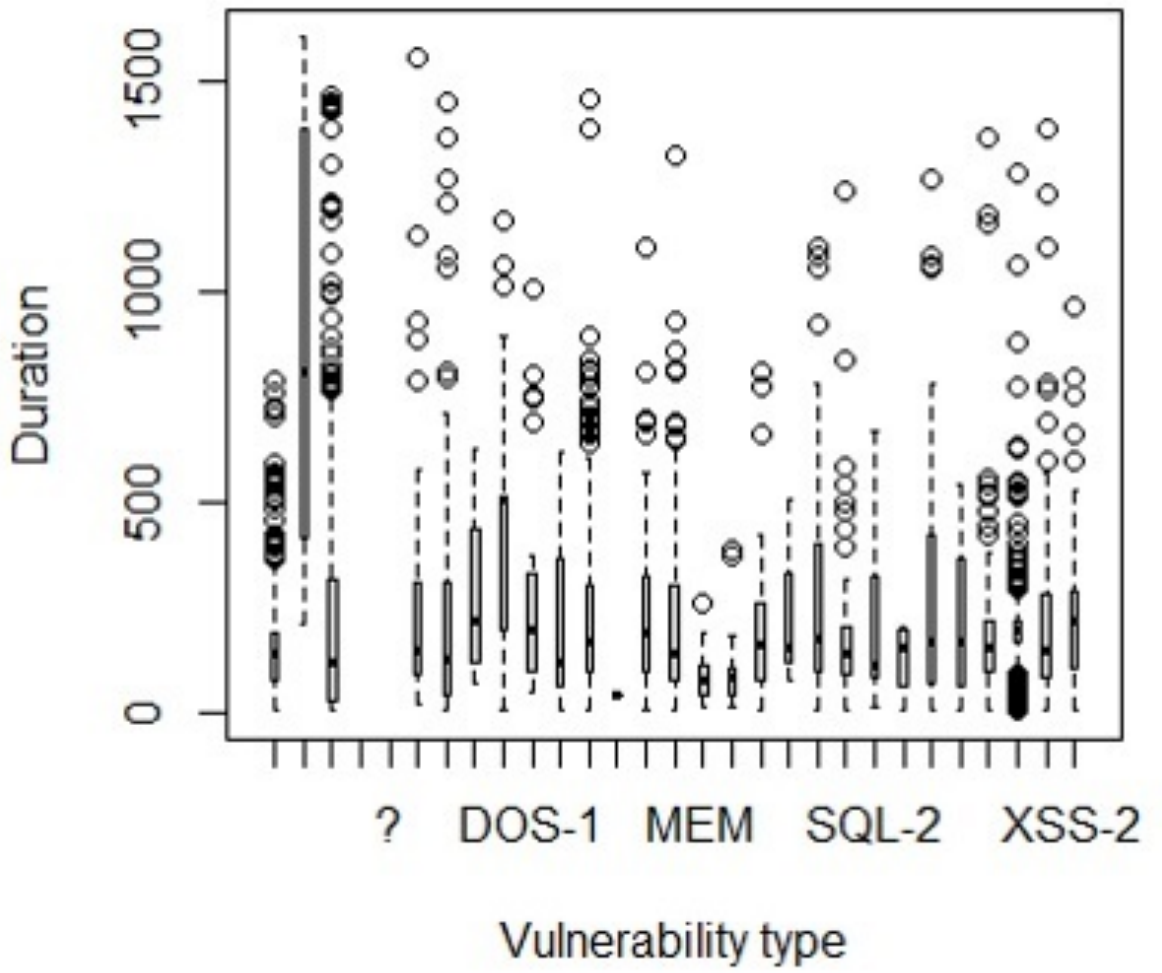
4. Descriptive components data
5. Descriptive projects data

Research Method



Data Exploration

Understand the data



Regression Methods

1. Linear Regression (LR)
2. Tree-based regression (RPART)
3. Neural-networks regression (NN)

Prediction Accuracy Metrics

1. Coefficient of determination
2. Prediction level (PRED) <= 25% error
3. Akaike Information Criterion – rate of info. Loss

Prediction Models (Parts)

Message source	Coef.
(Intercept)	249.17
Code scan tool	-50.04
Central security department	-38.05
Customers	-60.68
Ext. research organizations	-102.78
Int. development departments	-12.21
Test services	-124.74
Validation services	-21.88

84) vulnerabilitytype=,&OTHER,ACI-1,CDR-1,INF-1,MAC-1,MEM,XSS,XSS-2 270 5063771.00 286.53700

168) Component=AP-RC-ANA-UI-XLS,BC-BSP,BC-CST-DP,BC-CST-IC,BC-CTS-SDM,BC-CTS-TMS,BC-DOC-HLP,BC-DOC-TTL,BC-I18,BC-JAS-ADM-MON,BC-JAS-DPL,BC-SEC,BC-SEC-DIR,BC-SRV-ARL,BC-SRV-FSI,BC-UPG-SLM,BC-UPG-TLS-TLJ,BC-WD-CMP-FPM,BC-XI-CON-AXS,BC-XI-IBD,BC-XI-IBF,BI-BIP-AUT,BI-OD-STW,BI-RA-WBI,BW-BEX-OT-MDX,CA-GTF-IC-BRO,CA-GTF-IC-SCR,CA-GTF-RCM,CRM-BF,CRM-BF-SVY,CRM-CIC,CRM-IC-EMS,CRM-IC-FRW,CRM-IPS-BTX-APL,CRM-ISA,CRM-ISA-AUC,CRM-ISE,CRM-LAM-BF,CRM-MD-PRO,CRM-MKT-DAM,CRM-MKT-MPL,CRM-MSA,FS-CM,FS-SR,IS-A-DP,IS-U-CS-ISS,LO-AB-BSP,LO-GT,MFG-ME,MOB-APP-EMR-AND,PA-GE,PLM-PPM-PDN,PLM-WUI-RCP,PSM-GPR-SN,SBO-INT-B1ISN,SCM-EWM-RF,XAP-IC-IDM,XX-PROJ-CDP-TEST-296 119 1015233.00 205.82350 *

169) Component=AP-CFG,AP-LM-MON-HC,AP-LM-SUP,AP-RC-ANA-RT-MDA,AP-RC-RSP,AP-RC-UIF-RT,AP-SDM-EXC,BC-CCM-MON-OS,BC-CCM-SLD-JAV,BC-CST,BC-CUS-TOL-CST,BC-DB-ORA-INS,BC-DOC-TER,BC-ESI-WS-ABA,BC-ESI-WS-JAV-RT,BC-FES-BUS-RUN,BC-JAS-ADM-ADM,BC-JAS-COR,BC-JAS-SEC-UME,BC-MID-RFC,BC-SEC-SAL,BC-SRV-COM,BC-SRV-COM-FTP,BC-SRV-KPR-CS,BC-SRV-MCM,BC-SRV-SSF,BC-WD-ABA,BC-WD-

Accuracy of Prediction

DS \ Metric	Residual	AIC	PRED
ABAP	LR(0.526)	LR (122465)	LR (31.81%)
C ++ Java Cov./Fort.	LR (0.461)	LR (334565)	NN (33.81%)
Sec msg	LR (0.944)	RPART(6507)	RPART (34.71%)

Accuracy of Prediction with Extended Datasets

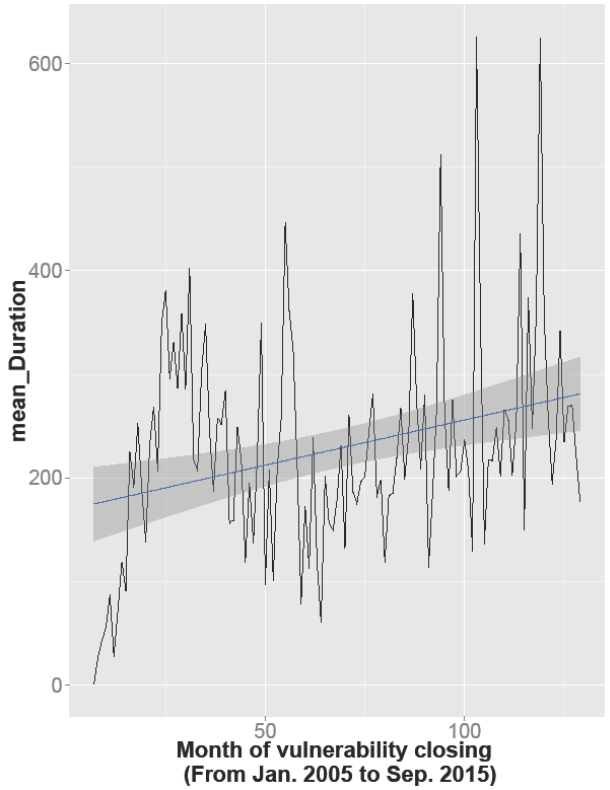
Metric / DS	Residual	AIC	PRED
ABAP	LR(0.526)	LR (122465)	LR (31.81%)
C ++ Java Cov./Fort.	LR (0.461)	LR (334565)	NN (33.81%)
Ext. C ++ Java	LR (1)	RPART(463)	LR (100%)
Sec msg	LR (0.944)	RPART(6507)	RPART (34.71%)
Ext. sec msg	LR (0.909)	RPART(6421)	NN (65.05%)

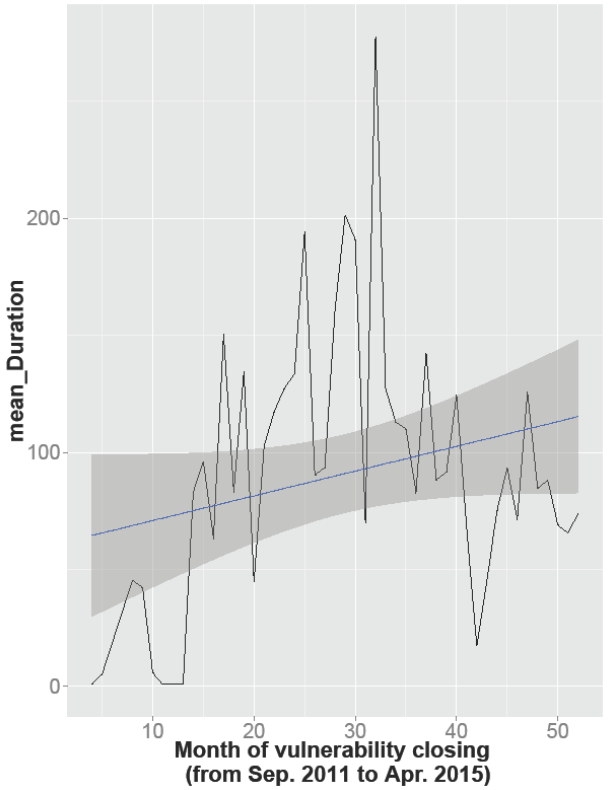
Can we use these methods?

Metric / DS	Residual	AIC	PRED
ABAP	LR(0.526)	LR (122465)	LR (31.81%)
C ++ Java Cov./Fort.	LR (0.461)	LR (334565)	NN (33.81%)
Ext. C ++ Java	LR (1)	RPART(463)	LR (100%)
Sec msg	LR (0.944)	RPART(6507)	RPART (34.71%)
Ext. sec msg	LR (0.909)	RPART(6421)	NN (65.05%)

What is wrong?

Sec. messages

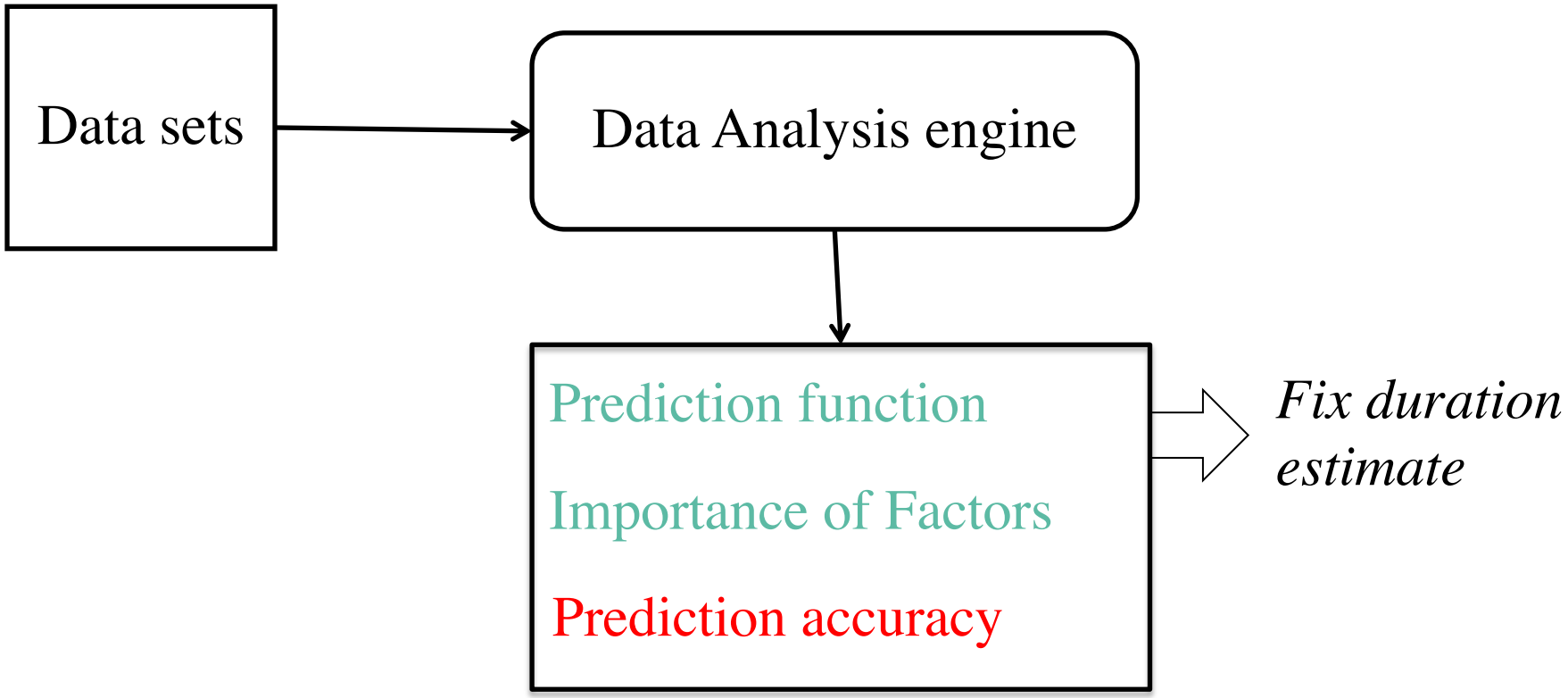




Fort./Cov.

22

The Prototype



1. Question: Is there a more accurate estimation method?

2. What is the Method?

3. What is the evidence?
4. Is that good enough?

Open Access | Published: 27 September 2016

Time for Addressing Software Security Issues: Prediction Models and Impacting Factors

Lotfi Ben Othmane ✉, Golriz Chehrazi, Eric Bodden, Petar Tsalovski & Achim D. Brucker

Data Science and Engineering **2**, 107–124(2017) | Cite this article

4497 Accesses | **4** Citations | **6** Altmetric | Metrics

Abstract

Finding and fixing software vulnerabilities have become a major struggle for most software development companies. While generally without alternative, such fixing efforts are a major cost factor, which is why companies have a vital interest in focusing their secure software development activities such that they obtain an optimal return on this investment. We investigate, in this paper, quantitatively the major factors that impact the time it takes to fix a given security issue based on data collected automatically within SAP's secure development

Link: https://link.springer.com/article/10.1007/s41019-016-0019-8

Thank you