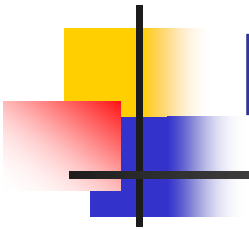
Who should fix this bug?

John Anvik, Lyndon Hiew, and Gail C. Murphy

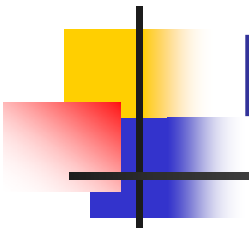
Presenter: Tao Xia

Sep 21, 2006



Motivition

- Triager: the person who go through bug reports and assign them to developers
- Propose a semi-auto process which
  - Can help triager to assigning bugs more efficient
  - New triager can work with a little knowledge of the organization



Processes (1)

- Characterizing bug reports
  - Use one-line summary and full text description
  - Remove all stop words, then use the remaining words to build a feature vector (used stemming)
- Labelling bug reports
  - It is project specific

Processes (2)

- Selecting Training Reports
  - Filtering the reports for training

Table 2: The effect of developer profile filtering on recommender accuracy and recall.

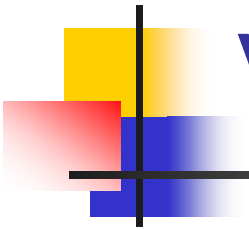
		# Dev.		Precision/Recall (%)	
		Firefox	Eclipse	Firefox	Eclipse
No Profile		414	146	23/1	58/7
>1 Fix in 3 mo.		94	82	59/2	57/7
Avg. Fixes Per Month Over 3 mo.	1	66	50	64/2	57/7
	2	33	42	59/2	57/7
	3	26	40	64/2	57/7
	4	21	40	64/2	58/7
	5	18	39	59/2	59/7
	6	13	37	45/1	57/7

Processes (3)

- Applying a machine learning algorithm
  - Support Vector Machines (SVM)
  - Naïve Bayes

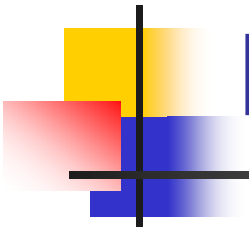
Table 3: The effect of different machine learning algorithms on recommender accuracy and recall.

Predictions	Naïve Bayes		SVM		C4.5	
	Firefox	Eclipse	Firefox	Eclipse	Firefox	Eclipse
1	59/2	54/6	64/2	58/7	64/2	40/5
2	59/2	49/11	52/3	52/13	41/3	34/9
3	59/2	44/15	57/6	47/16	42/5	31/12



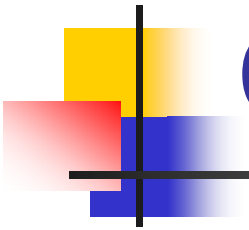
Validating the approach

- Validating the processes on GCC project
- Precision only 6%
- The processes do not work for all projects
  - One developer dominate the resolution activity
  - Labelling heuristic may not be accurate
  - The spread of bug resolution activity is low



Extension works

- Unsupervised machine learning
- Incremental machine learning
- Incorporating additional sources of data
- A component-based classifier



Conclusion and Future Work

- For Firefox and Eclipse, they achieved +50% precision
- The study shows the improvement of the bug assignment processes
- An empirical study of the approach
- Correlate additional information