

Traceability and Requirements Evolution

Group 2

Agenda

- Relation with the main topic
- Paper 1: Requirement Tracing on target (RETRO)
- Paper 2: A machine learning approach for tracing regulatory codes to product specific requirements
- Conclusions

Relation with the main topic

- Main papers talked about traceability in general
- The specific papers discuss different approach for doing traceability.
 - One using a tool call RETRO
 - The other approach is based on machine learning.



RTM and RETRO

RTM:

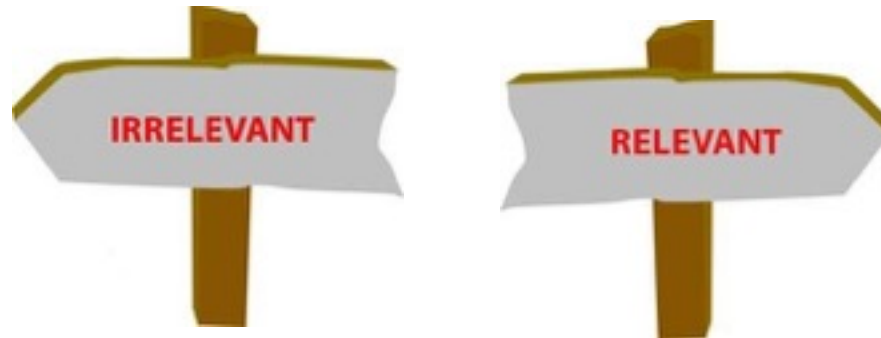
- Software maintenance require an up-to-date requirements traceability matrix (RTM)
- The generation and maintenance of RTM are tedious and error-prone

RETRO:

- Requirement Tracing on target (RETRO) is a tool that automates the generation of RTM
- RETRO helps to find more significant links
- RETRO address recovery of traceability using information retrieval (IR) and text mining to construct candidates trace

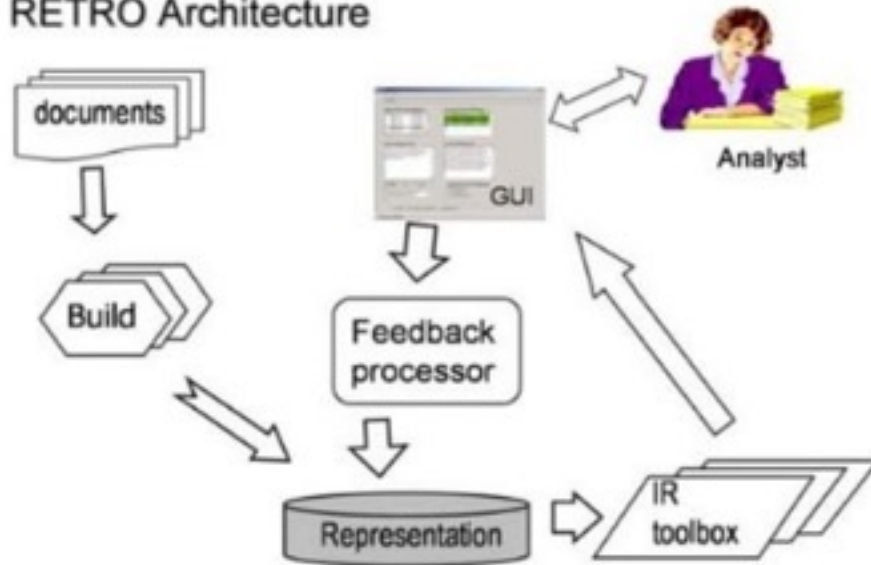
Information Retrieval (IR)

- IR involves a document collection and a user information need
- The information required by the user is represented in form of a query
- IR quality of retrieval can be improved by using user relevant feedback
- Relevance or irrelevance provided by human can be send back to IR system
- RETRO allows adjusting keyword weight of query vectors according to feedback



Architecture of RETRO

RETRO Architecture



RETRO 2.5 Functionality

- **Automatics tracing mode** (default method)
- **Manual tracing/browsing mode**: analyst browse low and high level documents to discover links not found by tool
- Trace all-at-one/one element at-a-time, feedback
- **Positive feedback**: belongs to final RTM
- **Negative feedback**: exclude link from RTM
- **Default**: analyst not ready to provide feedback.
- **Freezing high-level elements**: avoid retracing high-level requirements.
- **Browser tab functionality**: analyst can browse low-level and high-level requirements, if error or omissions are discovered, add newly discovered low-level element to RTM.

Interface of RETRO



Machine learning approach for tracing regulatory codes to product specific requirements

- Regulatory standards govern numerous areas of software intensive systems
- Employees must demonstrate that system meets all relevant regulatory codes
- The techniques are evaluated against security regulation from USA government's Health Insurance privacy and portability Act (HIPAA), traced against 10 healthcare products related requirement specifications



audit



Refactoring

Basic automatic tracing

- Probabilistic network model was used to generate traces from 10 targeted HIPAA regulation to requirements in each of the 10 patient healthcare systems.
- Automated methods have limited success for tracing regulatory codes
- Trace queries were formulated directly from the text of each HIPAA security regulation

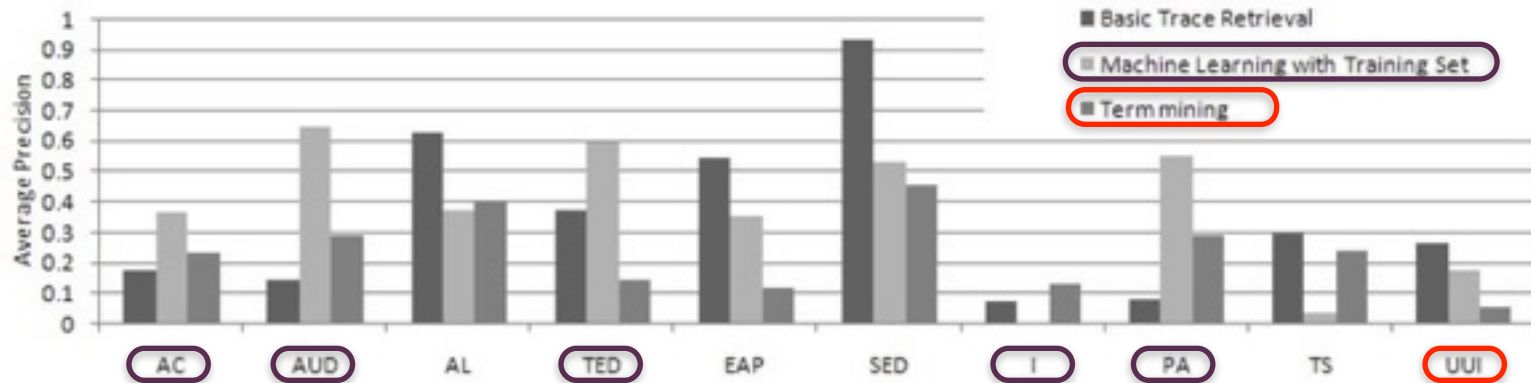
Machine learning approach

- The trace classification process includes the three primary stages of preparation, training, and classifying.
- Manually create traces matrices were used as both training and testing set.
- In each iteration, indicator terms were extracted from the requirements in the training set, and were weighted.
- The top 10 scoring indicators term for each regulatory code were selected and used to classify requirements.
- This technique is only feasible when training set is available

Web mining approach

- Initial queries for selecting domain specific documents were formulated by extracting heading of regulation. Researcher evaluate top results in BING, GOOGLE, and yahoo
- The process allows to find new terms and also allow to omit non-useful terms such as *persons* and *software* that occur in the original regulation
- Results shows that not all the queries performed well. Actually for a query the performance was worse than the last two processes. This is due to the risk of depending on finding a suitable an expressive set of documents

Approach results comparison



Benefit of having large training sets

problem due to not find suitable
set of documents

Conclusions

- Techniques described offer promising new approach for tracing regulatory codes
- Techniques can significantly reduce human effort and increase accuracy
- Techniques helps in maintaining and modifying the requirements

Questions?

