Enhancing IBM Requisite Pro with IR-based Traceability Recovery Features

Andrea De Lucia, Raffaele Landi, Rocco Oliveto, Genny Tortora
Department of Mathematics and Informatics, University of Salerno
84084, Fisciano (SA), Italy
adelucia@unisa.it, raffaele.landi81@gmail.com, {roliveto, tortora}@unisa.it

Workshop on Eclipse Technologies (Eclipse-IT 2008), Bari, Italy – November 18, 2008
Outline

• Background
  – Traceability recovery
  – Guidelines for the design of traceability recovery tools derived from empirical studies
  – Motivation: facilitate the technology transfer

• ReqTracer Pro: traceability recovery in IBM Requisite Pro
  – Architecture and functionalities

• Conclusion and future work
Background

Enhancing IBM Requisite Pro with IR-based Traceability Recovery Features

by Andrea De Lucia, Raffaele Landi, Rocco Oliveto, Genny Tortora

Workshop on Eclipse Technologies (Eclipse-IT 2008), Bari, Italy – November 18, 2008
Context

• Traceability...
  – the ability to describe and follow the artefact life-cycle
  – Example: a use case is implemented by one or more classes that are tested by
    a set of test cases

• Maintaining traceability links during software evolution
  – Tedious and error prone task
  – Often this information becomes out of date or it is completely absent

• Need for automatic support
  – Several methods have been proposed
    • Promising results achieved by IR-based tool
IR-based traceability recovery

 Ranked list
 Candidate links

<table>
<thead>
<tr>
<th>Source</th>
<th>Target</th>
<th>Recovery</th>
<th>Source</th>
<th>Target</th>
<th>Recovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source_1</td>
<td>Target_2</td>
<td>95.4%</td>
<td>Source_3</td>
<td>Target_4</td>
<td>92.1%</td>
</tr>
<tr>
<td>Source_1</td>
<td>Target_1</td>
<td>85.6%</td>
<td>Source_2</td>
<td>Target_1</td>
<td>83.2%</td>
</tr>
<tr>
<td>Source_2</td>
<td>Target_2</td>
<td>83.2%</td>
<td>Source_3</td>
<td>Target_3</td>
<td>81.2%</td>
</tr>
<tr>
<td>Source_3</td>
<td>Target_3</td>
<td>79.0%</td>
<td>Source_1</td>
<td>Target_3</td>
<td>77.5%</td>
</tr>
<tr>
<td>Source_2</td>
<td>Target_4</td>
<td>64.3%</td>
<td>Source_3</td>
<td>Target_3</td>
<td>53.2%</td>
</tr>
<tr>
<td>Source_3</td>
<td>Target_1</td>
<td>43.9%</td>
<td>Source_3</td>
<td>Target_1</td>
<td>38.7%</td>
</tr>
<tr>
<td>Source_1</td>
<td>Target_4</td>
<td>23.6%</td>
<td>Source_2</td>
<td>Target_1</td>
<td>20.7%</td>
</tr>
</tbody>
</table>

Source artefacts

IR-based Traceability Recovery Tool

Target artefacts
The Precision/Recall problem

• Two metrics to measure the performances of IR-based tools
  – Recall = (|correct ∩ retrieved|) / |correct|
  – Precision = (|correct ∩ retrieved|) / |retrieved|
  – Low precision => high number of false positives to discard

• Recovering all correct links is in general impractical!
  – Necessary to use a very low threshold
  – Low threshold => High number of links retrieved => Low precision
  – High effort to discard too many false positives

• Results from a case study show that...
  – about 50,000 false positives have to be discarded by the software engineer in order to trace 361 links among about 200 artefacts
Density of correct links and false positives

- Automatic tracing should be combined with manual tracing!
Incremental traceability recovery

The software engineer decides to stop the process as the effort to discard false positives is becoming too high. Probably he does not retrieve all correct links!
Motivations

- Lacks of IR-based traceability recovery tools
  - In general, stand-alone tools
- A traceability recovery tool is needed
  - Case studies and controlled experiments were carried out
  - The tool drastically reduce the time to complete a traceability recovery task and reduce tracing errors
  - Generally subjects traced about 60-70% of the links with the tool support
- Facilitate the technology transfer of traceability recovery tools
  - Its usefulness should be investigated through empirical user studies
  - It should be integrated within a commercial and widely used CASE tool
ReqTracer Pro
Traceability recovery in IBM Requisite Pro

Enhancing IBM Requisite Pro with IR-based Traceability Recovery Features
by Andrea De Lucia, Raffaele Landi, Rocco Oliveto, Genny Tortora

Workshop on Eclipse Technologies (Eclipse-IT 2008), Bari, Italy – November 18, 2008
IBM Requisite Pro and Modeler

• IBM Requisite Pro
  – Requirements management tool
  – It provides support for traceability, but
  – Traceability information has to be manually managed

• IBM Modeler
  – A collaborative platform for visual modeling and design
  – It integrates several products of IBM Rational Suite, including Requisite Pro
  – It is built on top of the Eclipse platform

• Why such a choice...
  – Traceability links are stored in Requisite Pro
  – IBM Modeler provides a high level of extensibility
  – It also provides API allowing the communication with Requisite Pro
  – A plug-in for Modeler (Eclipse) has the same UI of Requisite Pro
ReqTracer Pro: recovery process

TERM TRANSFORMATION EXAMPLES
- Composed Term separation
- Lower Case letter reduction
- Stop word removal
- Stemming
- ...

Source Artefact Set

Term extraction

Target Artefact Set

Term extraction

Indexer

Document classifier

Candidate links

Ranked list cutter

Cut threshold
ReqTracer Pro: classifier

- A traceability recovery tool based on Vector Space Model
  - VSM represents artefacts as vector of terms (extracted from the artefacts)
  - The similarity between two artefacts is represented by the cosine of the angle between the two artefact vectors
ReqTracer Pro: architecture
ReqTracer Pro: traceability recovery
ReqTracer Pro: link analysis

- A list or a matrix of candidate links is provided
ReqTracer Pro: search engine

- A search engine is also integrated in Requisite Pro
Conclusion and future work

Enhancing IBM Requisite Pro with IR-based Traceability Recovery Features

by Andrea De Lucia, Raffaele Landi, Rocco Oliveto, Genny Tortora
Conclusion

• ReqTracer Pro
  – An IR-based traceability recovery tool integrated in Requisite Pro

• The design of the tool is based from empirical studies

• Why Requisite Pro?
  – Facilitate the technology transfer

• Future work
  – Further experimentation
  – Improvement of the IR engine
Thank you!

Questions and/or comments

Andrea De Lucia, Raffaele Landi, Rocco Oliveto, Genny Tortora
Department of Mathematics and Informatics, University of Salerno
84084, Fisciano (SA), Italy
adelucia@unisa.it, raffaele.landi81@gmail.com, {roliveto, tortora}@unisa.it

Workshop on Eclipse Technologies (Eclipse-IT 2008), Bari, Italy – November 18, 2008