**Name of the Paper:** Textual Requirement Analysis for UML Diagram Extraction by using NLP

S.D.Joshi, PhD. BVCOE, Pune, Maharashtra, India.

Dhanraj Deshpande BVCOE, Pune, Maharashtra, India.

International Journal of Computer Applications (0975 – 8887)

Volume 50 – No.8, July 2012

**Methodology discussed in the paper:** Requirement Analysis and UML Diagram Extraction (RAUE).

**Existing tools mentioned and their shortcomings:**

* **CIRCE**

CIRCE is simple to operate but lacks efficiency in generating UML models from the analysed requirements.

* **LIDA (Linguistic Assistant for Domain Analysis)**

LIDA just identifies the list of nouns, adjectives, verbs and the user has to decide which word goes in classes or attributes or operations

LIDA depends on the knowledge of the user

If the user has less knowledge, then this tool may ignore some important functionality of the system.

* **UCDA (Use Case Driven Development Tool for class model generation)**

The software designer has to analyse requirement and identify the objects by Object Model Creation Process.

* **CM-Builder**

CM-Builder cannot identify operations for candidate classes.

**RAUE Algorithm:**

Step 1) Textual document is given as a input document to RAUE filtering model.

Step 2) Process the text document & tokenize the user requirement document.

e.g. Library contains books and journals. This is input sentence which is afterword tokenized & produces output as, <Library>, <Contains>,<Books>,<and>,<journals>,<.>

Step 3) searching of stop words & result of stop words is stored in “vocabulary-of-stopwords” list by using Morphological analysis.

Step 4) Defines total number of word in text document by removing stop word.

Step 5) Find out total occurrences of concept word in document by morphological & POS Tagging analysis.

Step 6) Use OpenNLP parser for segmentation, POS tagging, chunking of the textual document. It also does lexical & syntactical Parsing of document.

Step 7) In this step, result of open NLP parser is stored in “vocabulary-of-concepts” list. Result contain noun, verb, helping verb, adjective, article & noun phrase (NP).

e.g. Library contains books.

Step 8) Use of Step 2 to Step 7 for finding out “vocabulary-of- stopwords” and “vocabulary-of-concepts”.

Step 9) In this step, different relationship between concepts words from “vocabulary-of-concepts” is recognized. It can extract different relationship like generalization, association, aggregation and dependency. It also finds out one-to-one, one-to-many and many-to-many.

Step 10) At this step, UML diagram model is extracted by using previous steps information.

Step 11) Final step in which user can interact with diagram and make changes if required and save it.

RAUE gives result of 85% in terms of recall value and precision value.