# Summary Report (08/14/2014)

## Overview

In this week, I have done the following things:

## Shallow Summarization with TF-IDF

Instead of using the unigram frequency, I tried to use the TF-IDF weight. Actually, the frequency part is TF. It is just a normalization by including IDF.

 \mathrm{idf}(t, D) =  \log \frac{N}{|\{d \in D: t \in d\}|}

I used two different ways to compute the IDF by assuming.

* All the documents are belong to one cluster (tfidf1)
* The documents are divided into three clusters (POI, MP, LP) (tfidf3)

It will affect both the numerator and denominator to get the IDF.

### Results

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | POI | | | MP | | | LP | | |
|  | R1 | R2 | R-SU4 | R1 | R2 | R-SU4 | R1 | R2 | R-SU4 |
| Unigram | 35.22% | 1.72% | 9.19% | 34.61% | 1.95% | 8.71% | 23.55% | 0.89% | 4.22% |
| TopicS-nostemming | 30.42% | 1.18% | 7.36% | 33.81% | 0.72% | 8.74% | 17.87% | 0.37% | 2.66% |
| TopicS-stemming | 29.98% | 0.92% | 7.04% | 30.91% | 0.27% | 7.16% | 19.14% | 0.37% | 2.88% |
| NP-Hard | 26.98% | 4.30% | 5.85% | 27.75% | 5.77% | 6.44% | 18.11% | 0.16% | 2.67% |
| NP-Soft | 35.83% | 8.74% | 9.95% | 34.11% | 6.92% | 8.74% | 22.43% | 1.93% | 4.00% |
| Unigram-tfidf1 | 30.14% | 2.15% | 7.69% | 31.37% | 1.07% | 7.40% | 13.12% | 0.41% | 1.62% |
| Unigram-tfidf3 | 31.05% | 0.40% | 7.77% | 32.18% | 1.62% | 7.70% | 10.47% | 0.00% | 1.03% |

### Observations:

* Like (Yatani et al, 2011) [2], the TF-IDF weighting doesn’t get better result than TF. Probably because the common words are also useful, such as activity, group, class, etc.

## NP Summarization with semantic similarity

By using the SEMILAR toolkit (http://deeptutor2.memphis.edu/Semilar-Web/public/semilar-api.html), it is able to compute semantic similarity in a variety of ways.

Now, I have successfully make the toolkit to work and to compute Similarity between two phrases, such as greedyComparerWNLin(various materials, how materials) = 0.5, which uses use wordnet LIN method for Word 2 Word similarity.

In this way, I can get pair-wise similarity between two phrases. The problem is how to get individual weight for a single phrase. Now, I’m taking an average method.

### Results:



### Observations:

* The WordNet\_Lin similarity method is better than the hard frequency (both R1 and R2)
* However, this semantic metric is still not better than the NP-Soft model

## TODO

In the coming days, I will try different semantic metrics given by SEMILAR, including LSA, BLEU, etc.

## Paper I read

[1] Yatani, Koji, et al. "Review spotlight: a user interface for summarizing user-generated reviews using adjective-noun word pairs." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2011. [<http://nb.mit.edu/f/15423>]

[2] Yatani, Koji, et al. "Analysis of adjective-noun word pair extraction methods for online review summarization." *IJCAI Proceedings-International Joint Conference on Artificial Intelligence*. Vol. 22. No. 3. 2011. [<http://nb.mit.edu/f/15424>]