

## International Journal of Computer Science and Mobile Computing



A Monthly Journal of Computer Science and Information Technology

ISSN 2320-088X

*IJCSMC, Vol. 3, Issue. 2, February 2014, pg.382 – 388*

### RESEARCH ARTICLE

# Improvement of Email Summarization Using Statistical Based Method

**Mithak I. Hashem**

Department of Computer Science, Thi-Qar University, Iraq

Mithakh@yahoo.com

---

**Abstract** - *Automatic text summarization is undergoing wide research and gaining importance as the availability of online information is increasing. Email is one of the most important online tools that many of us depends on in his everyday life. Finding Email summaries may be crucial for many users. We deal with email text as a single-document in this research. Text summarization can be classified into two approaches: extraction and abstraction. This research focuses on extractive one. The goal of text summarization based on extraction approach is sentence selection. Our proposed method to obtain the suitable sentences is to assign some numerical measure of a sentence (statistically) for the summary called sentence score and then select the best ones to be included within Email summary. The most important step in summarization by extraction is the identification of important features. In our experiment, we used 130 test Email text from Enron\_Sent-Mail\_Sample data set. Each Email document is prepared by preprocessing process: sentence segmentation, tokenization, removing stop word, and word stemming. Then, we used 7 important features and calculate their score for each sentence. The results show that the best average similarities with the reference summary (gold summary) were obtained by our method.*

**Keywords-** *email summarization; single-document summarization; sentence feature; sentence score; similarity*

---

Full Text: <http://www.ijcsmc.com/docs/papers/February2014/V3I2201488.pdf>