

# **Full-Text Search in Myanmar Language using Apache Lucene**

**Nang Zin Min Aye, Khin Mar Soe**

*University of Computer Studies, Yangon*

*nangzinminaye@ucsy.edu.mm,*

*khinmarsoe@ucsy.edu.mm*

## **Abstract**

*In text retrieval, full-text search refers to techniques for searching a single computer-stored document or a collection in a full-text database. However, when the number of documents to search is potentially large, or the quantity of search queries to perform is substantial, the problem of full-text search is often divided into two tasks: indexing and searching. This paper proposes a full text search in Myanmar documents stored in text and printed document form format. The indexing is based on inverted structure and the searching is done using Apache Lucene engine. Although there are many systems support file content search, searching contents in Myanmar language is not widely used. So, it is found that the proposed system is useful for the content search in files such as PDFs, Microsoft Word and text files written in Myanmar language.*