

# CONTINUOUS TOP-K MONITORING ON DOCUMENT STREAMS

*A project report submitted in partial fulfillment of the requirements for  
the award of the Degree of*

**BACHELOR OF TECHNOLOGY**

In

**COMPUTER SCIENCE AND ENGINEERING**

Submitted by;

R.H.V.HARSHINI Regd. No.15811A0564 M.P.SAI RAMAYYA Regd.No.16815A0506

S.RAJESH Regd.No.15811A0566 P.MOUNIKA Redg.No.12811A0563

G.V.D.SAIRAM Regd.No.16815A0503

Under the guidance of

V.TRINADHA

*Assistant professor*

*Department of Computer Science and Engineering*



**AVANTHI INSTITUTE OF ENGINEERING & TECHNOLOGY**

(Approved by AICTE, New Delhi & Permanently Affiliated to JNTU Kakinada) (Accredited by NAAC,  
UGC & NBA, AICTE) MAKAVARAPALEM, NARSIPATNAM,

VISAKHAPATNAM DIST (2105-2019)

**AVANTHI INSTITUTE OF ENGINEERING & TECHNOLOGY**

(Approved by AICTE, New Delhi & permanently affiliated to JNTU Kakinada)

(Accredited by NAAC, UGC & NBA, AICTE)

TAAMARAM, MAKAVARAPALEM, NARSIPATNAM - 531113

VISAKHAPATNAM (DIST)



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**CERTIFICATE**

This is to certify that the project entitled "CONTINUOUS TOP-K MONITORING ON DOCUMENT STREAMS" in partial fulfillment for the of degree of Bachelor of Technology in COMPUTER SCIENCE AND ENGINEERING, at AVANTHI INSTITUTE OF ENGINEERING AND TECHNOLOGY, MAKAVARAPALEM, VISAKHAPATNAM is an bonafide work carried out by R.H.V.HARSHINI (15811A0564), M.P.SAIRAMAYYA (16815A0506), S.RAJESH(15811A0566), P.MOUNIKA(15811A0506), G.V.D.SAIRAM (16815A0503), under the guidance and supervision during 2018-2019.

Project Guide

External Examiner

Head of the Department

## ABSTRACT

The efficient processing of stream plays an important role in many information filtering systems. Emerging applications, such as news update filtering and social network notifications, demand presenting end-users with the most relevant content to their preferences. In this work, user preferences are indicated by a set of keywords. A central server monitors the document stream and continuously reports to each user the top-k documents that are most relevant to her keywords. Our objective is to support large numbers of users and high stream rates, while refreshing the top-k results almost instantaneously. Our solution abandons the traditional frequency-ordered indexing approach. Instead, it follows an identifier-ordering paradigm that suits better the nature of the problem. When complemented with a novel, locally adaptive technique, our method offers

(i) Proven optimality the number of considered queries per stream event, and

(ii) An order of magnitude shorter response time (i.e., time to refresh the query results) than the current state-of-the-art.