



**RESEARCH ARTICLE**

# Grouping and Categorization of Documents in Relativity Measure

*V.Asaiithambi<sup>1</sup>, D.John Aravindhar<sup>2</sup>, V.Dheepa<sup>3</sup>*

<sup>1</sup>PG student of Computer Science and Engineering, Hindustan University, Padur, Chennai, India

<sup>2</sup>Assistant Professor, Department of Computer Science and Engineering, Hindustan University,  
Padur, Chennai, India

<sup>3</sup>Assistant Professor, Department of Computer Science and Engineering, Hindustan University,  
Padur, Chennai, India

<sup>1</sup>asaivenkat@gmail.com; <sup>2</sup>jaravindhar@hindustanuniv.ac.in; <sup>3</sup>vdeepa@hindustanuniv.ac.in

---

***Abstract— This paper presents a spectral clustering method called correlation through preserving indexing (CPI), which is to perform in the correlation similarity measure space. The documents are considered into a low dimensional semantic space, the correlations between the documents in the local patches are maximized and correlations between the documents outside these patches are minimized. The intrinsic structure of the document space is included in the similarities between the documents. Correlation is the similarity measure for finding the intrinsic structure of the document space than Euclidean distance. Simultaneously, the proposed CPI methods can effectively finding the intrinsic structures included in high-dimensional document space. The effectiveness of the new method is implemented by extensive experiments conducted on various data sets and by comparison with existing document clustering methods.***

***Key Terms: - Document Clustering, Correlation Latent Semantic Indexing, Dimensionality Reduction, Correlation Measure.***

---

Full Text: <http://www.ijcsmc.com/docs/papers/March2013/V2I3201321.pdf>