IMPACT: International Journal of Research in Engineering & Technology (IMPACT: IJRET)
ISSN(E): 2321-8843; ISSN(P): 2347-4599
Vol. 2, Issue 4, Apr 2014, 129-136

CLUSTERING FOR FORENSIC ANALYSIS

SAHIL KHENAT, PRATIK KOLHATKAR, SARANG PARIT & SHARDUL JOSHI

Department of Computer Engineering, UCOER, Pune University, Pune, Maharashtra, India

ABSTRACT

In today's digital world, information in computers has great importance and this information is very essential in context for future references and studies irrespective of various fields. So surveying of such information is critical and important task. In computer forensic analysis, a lot of information present in the digital devices (computers in context of our paper) is examined to extract information. And computer consists of hundreds of thousands of files which contain unstructured text or information, so clustering algorithms are of great interest. Clustering helps to improve analysis of documents under consideration. This document clustering analysis is very useful for crime investigations to analyze the information from seized digital devices like computers, laptops, hard disks, tablets. There are total six algorithms used for clustering of documents like K-means, K-medoids, single link, complete link, Average Link, CSPA. These six algorithms are very efficiently used to cluster the digital documents. These algorithms make the analysis very fast by clustering very close documents in one cluster. Also two validity index are used to find out how many clusters are formed.

KEYWORDS: Clustering, CSPA, K-Mean, K-Medoids