Abstract

A Text stream is a sequence of chronologically ordered documents, being generated in various forms. Multiple text streams that are correlated to each other by sharing common topics. Our aim is to extract the knowledge of the text stream from the listed documents. In particular, vulnerabilities could include compromise of data security and loss of information which leads to data leakage. To provide a data security and privacy a key management is used. Documents from different sequences about the same topic may have different time stamps termed as asynchronous. Here we first, use Apriori Algorithm to extract the common topics for the search text from the given data set based on the time stamps using Timestamp-Based Protocols. We also use formetric encryption algorithm, which combines Encryption and integrated key management to protect and control access to sensitive files on file servers. Second, Ranking is
involved in both admin side and user side of mining work which is based on usability of documents.

References

- Hyungsul Kim, Yizhou Sun, Julia Hockenmaier and Jiawei Han &quot;ETM: Entity topic models for mining documents associated with entities&quot; Data Mining(ICDM), 2012 IEEE 12th International conference on digital object, 2012
- Qiaozhu Mei &quot;Discovering Evolutionary Theme Patterns from Text An Exploration of Temporal Text Mining&quot; KDD &apos;05 Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining, 2005
- Thomas Hofmann International, Berkeley, &quot;Probabilistic Latent Semantic Indexing&quot; Proceedings of the Twenty- Second Annual International SIGIR Conference on Research and Development in Information Retrieval CA & EECS Department, CS Division, UC
- Xing Yi and James Allan &quot;Evaluating topic models for information retrieval&quot; CIKM &apos;08 Proceedings of the 17th ACM conference on Information and knowledge management, 2008.
- Xiaochuan Ni, Jian-Tao Sun, Jian Hu, Zheng Chen &quot;Mining Multilingual Topics from Wikipedia&quot; WWW &apos;09 Proceedings of the 18th international conference on World wide web, 2009.
Index Terms

Computer Science
Information Sciences

Keywords
Mining multiple text sequence  Ranking  key management.