



# Implementation On Unstructured Data Using Picture Key Cryptographic Mechanism

<sup>1</sup>Meena Kumari, <sup>2</sup>Mrs. Deepika

<sup>1</sup>Research Scholar, Deptt. Of Computer Engineering, IJET Kinana (Jind)

<sup>2</sup>Assistant Professor , Deptt. Of Computer Engineering, IJET Kinana (Jind)

Peacrfulboy8060@gmail.com, Deepika.ad14@gmail.com

**ABSTRACT:-** Unstructured data refers to information that does not had been a pre-defined data model or is not organized in a pre-defined manner. Unstructured information is typically text-heavy, but might contain data such as dates, numbers, & facts as well. This result in irregularities & ambiguity that make it difficult to understand using traditional programs as compared to data stored in fielded form in database in documents. Then we would convert unstructured data in form of matrix. In this research we have used MATLAB as simulation tool. Here we would make comparative analysis of traditional unstructured data security & proposed unstructured data.

**Keyword:** Data Warehouse, Data mining, MATLAB, Unstructured data, Data security, XOR, Cryptography

## 1. INTRODUCTION

The goals of an Unstructured Data Management System (UDMS) are difficult to outline because everyday researchers & other industry players find new problems & new applications.

The most well known example of search over unstructured data is Google [3], that provide to users a search tool about Internet (probably bigger source of unstructured information). Live plasma [4] is a good example of an application that shares these goals, showing relations among films or among musicians, in a innovative user interface. An UDMS should provide a rime work that support building applications within one or more of these goals.

One of major differences between structured & unstructured environments is that in structured environment data is updated on a regular basis. Whenever your ATM is used, whenever a deposit is made, whenever a check is cashed, your bank account is updated. & it is at these times that

structured data is commonly measured, or some record is created.

## 2. UNSTRUCTURED RESOURCE ANALYSIS

Unstructured data [5] had been no identifiable structure. Unstructured data typically includes bitmap images/objects, text & other data types that are not part of a database.

Unstructured data could be textual or non-textual. Textual unstructured data is generated in media [17] like email messages, PPT, DOC, collaboration software & instant messages. Non-textual unstructured data is generated in media like JPEG images, MP3audio files & Flash video files. Some current technologies used for content searches on unstructured data have need of tagging entities such as names or applying keywords & Meta tags.

## 3. CRYPTOGRAPHY

It had been discipline of information security had been called Cryptography. Meaning of Cryptography

ISSN : 2278-6848



© International Journal for  
Research Publication and Seminar



had been “hidden” imitative from Greek kryptos. Cryptography means hide information within storage or transfer including methods such as microdots, integration of words with image.

Cryptography had been process of altering plaintext (ordinary text, just as letter) using process encryption into cipher text using procedure decryption. This procedure had been used to secure communication between two parties within occurrence of third party. There are four goals for Modern cryptography:

### Confidentiality

It identifies that only participants (Sender & Receiver) should be able to access message.

### Integrity

Content of message should not be changed. If it had been altered, then it had been called type of modification attack.

### Non-repudiation

There had been situation where sender converts content of message & after that he refuses that he had not sent message.

### Authentication

Both sender & receiver had to prove credentials to each other.

In current times, cryptography had been basic requirement of computer experts for security purposes so that two parties could send data to each other without any modification & confidently. So both sender & receiver could validate to each other for secure communication so that material could be safely send to each other.

## SECURITY MECHANISMS

### Spread-Spectrum:-

Spread spectrum, first developed by military, became increasingly popular largely due to its

interference tolerance & coexistence capabilities. Today spread spectrum outside military-only realm, ranges from digital cellular phones & wireless PCs to wireless LANs.

**Prior Messaging:-** Instant messaging is a set of communication technologies used for text-based communication between two or more participants over Internet or other types of networks. Some systems permit messages to be sent to users not then logged on thus removing some differences between IM & email often done by sending message to associated email account.

## 4. RESULT AND DISCUSSION

### Decryption process:

To decrypt encrypted image encfinal.jpg we again run guide command for further process:

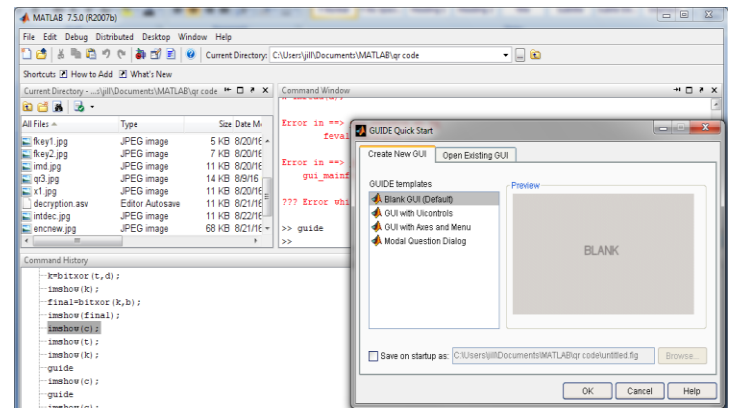


Fig 1 Decryption process

In order to decrypt unstructured data code we would open Decryption.fig in matlab using guide command

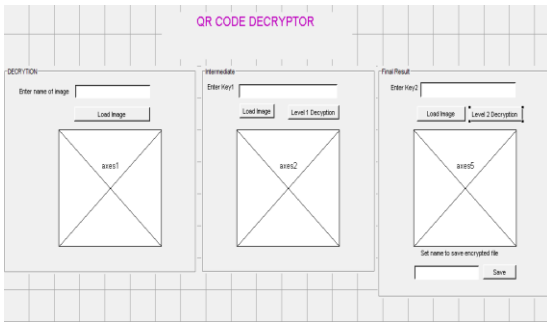


Fig 1 Unstructured data Code Decrypt or Select decryption.fig to load fig file

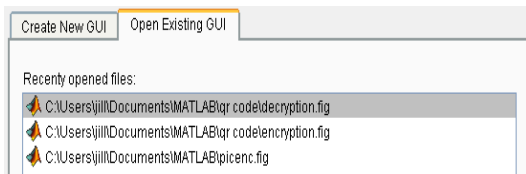


Fig 2 Select decryption.fig to load fig file

Encfinal.jpg image is loaded for decryption.



Fig 3 Encfinal.jpg image is loaded for decryption

Fkey2.jpg image is loaded for level 1 Decryption.

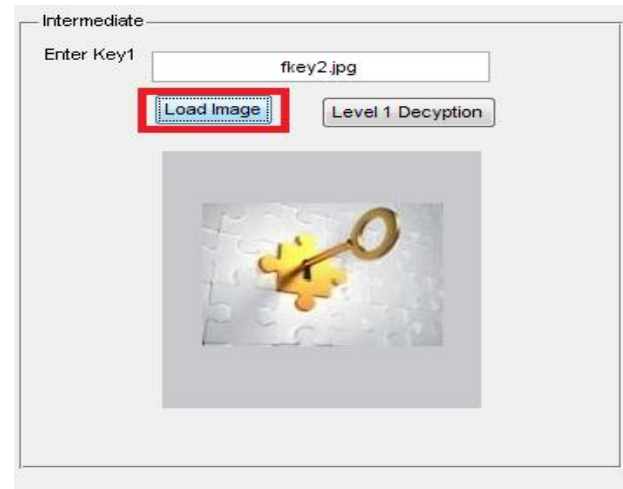


Fig 4 Fkey2.jpg image is loaded Level 1 Decryption image is shown here.



Fig 5 Level 1 Decryption image is shown

Fkey1.jpg image is loaded for level 2 Decryption.

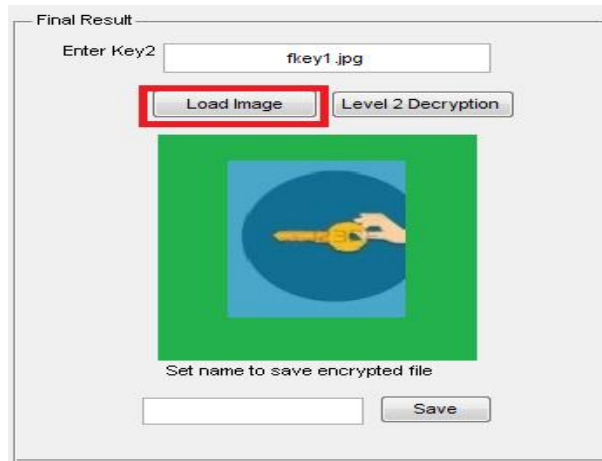


Fig 6 Fkey1.jpg image is loaded for level 2  
Decryption

Final level 2 Decrypted image.

## 5. CONCLUSION

Unstructured textual data is being constantly generated via call centre logs, emails & documents on web, blogs, tweets, customer comments, customer reviews, & so on. While amount of textual data is increasing swift, ability to make summary, understand, & make sense of such data for making better firm decisions pending challenging. Investigation of new challenges in data ware house management in unstructured data. We have done Convert unstructured data in form of matrix. We made secure unstructured data using bit xor operation in mailable have also made comparative analysis of traditional unstructured data security & proposed unstructured data. Since all enterprises like manufacturing, IT Tools & Services, Information Services, Government, Education, non-profit organization, Financial Insurance works within majority of data like unstructured data which have been increasing significantly every day & also it is big business asset & not liability, Hence analysis of

unstructured big data both textual & non-textual together called streaming or unstructured data in motion is a really recent challenge .This paper suggests answer that by providing awareness & education & also by establishing good governance policies & alternative technologies we could resolve this difficulty in easier way which is very important about unstructured large scale data . These innovative approaches to computation make analytics possible & make them affordable.

## REFERENCE

- [1] E. Agichtein, L. Gravano, J. Pavel, V. Sokolova, A. Voskoboynik. Snowball: A prototype system for extracting relations from large text collections. In SIGMOD, 2001.
- [2] S. Brin. Extracting patterns relations from world wide web. In WebDB, 1998.
- [3] Y. Cai, X. Dong, A. Y. Halevy, J. Liu, J. Madhavan. Personal information management with semex. In SIGMOD, 2005.
- [4] X. Chai, B. Vuong, A. Doan, J. F. Naughton. Efficiently incorporating user interaction into extraction integration programs. Technical Report UW-CSE-2008, University of Wisconsin-Madison, 2008.
- [5] F. Chen, A. Doan, J. Yang, R. Ramakrishnan. Efficient information extraction over evolving text data. In ICDE, 2008
- [6] E. Chu, A. Baid, T. Chen, A. Doan, J. F. Naughton. A relational approach to incrementally extracting querying structure in unstructured data. In VLDB, 2007.
- [7] P. DeRose, X. Chai, B. Gao, W. Shen, A. Doan, P. Bohannon, X. Zhu. Building community wikipedias: A machine-human partnership approach. In ICDE, 2008.



[8] P. DeRose, W. Shen, F. Chen, A. Doan, R. Ramakrishnan. Building structured web

community portals: A top-down, compositional, incremental approach. In VLDB, 2007.

[9] P. DeRose, W. Shen, F. Chen, Y. Lee, D. Burdick, A. Doan, R. Ramakrishnan. Dblife: A community information management platform for database research community (demo). In CIDR, 2007.

[10] A. Doan. Data integration research challenges in community information management systems, 2008. Keynote talk, Workshop on Information Integration Methods, Architectures, Systems (IIMAS) at ICDE-08.

[11] A. Doan, P. Bohannon, R. Ramakrishnan, X. Chai, P. DeRose, B. Gao, W. Shen. User-centric research challenges in community information management systems. IEEE Data Engineering Bulletin, 30(2):32–40, 2007.

[12] A. Doan, J. F. Naughton, A. Baid, X. Chai, F. Chen, T. Chen, E. Chu, P. DeRose, B. Gao, C. Gokhale, J. Huang, W. Shen, B. Vuong. case for a structured approach to managing unstructured data. In CIDR, 2009.

[13] A. Doan, R. Ramakrishnan, F. Chen, P. DeRose, Y. Lee, R. McCann, M. Sayyadian, W. Shen. Community information management. IEEE Data Engineering Bulletin, 29(1):64–72, 2006.

[14] A. Y. Halevy, M. J. Franklin, D. Maier. Principles of dataspace systems. In PODS, 2006.

[15] J. Huang, T. Chen, A. Doan, J. F. Naughton. On provenance of non-answers to queries over extracted data. PVLDB, 1(1):736–747, 2008.

[16] R. Krishnamurthy, Y. Li, S. Raghavan, F. Reiss, S. Vaithyanathan, H. Zhu. Systemt: A system for declarative information extraction, 2008. SIGMOD

Record, Special Issue on Managing Information Extraction.