EGU21-309
https://doi.org/10.5194/egusphere-egu21-309
EGU General Assembly 2021
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Data-In-Image: a novel concept for data manipulation and encryption implementation

Mohamed Abdelwahed[1,2]
[1]King Abdulaziz University, Geohazards Research Center, Seismic Hazards Research Unit, Jeddah, Saudi Arabia (mfibrahim@kau.edu.sa)
[2]National Research Institute of Astronomy and Geophysics (NRIAG), Cairo, Egypt (mfarouk40@yahoo.com)

Conventionally, the way of storing and exchange numerical data depends mainly on binary data files in compressible form. In this era of the Big Data and machine learning systems and with the accumulation of data with different forms and types, it is important to find an alternative way for handling the data. The binary data are software dependent which does not exhibit its content and type without accessing the data by the proper software. In addition, it does not have any encryption ability. To solve this issue, we propose a new concept to handle the digital data in a descriptive, encrypted, compressed form, and able to be previewed. The idea is to pack the binary bits into a bitmap image with specific coding scheme. This approach employs the Steim scheme as a primary compression tool with a 128-bit encryption method then packs the encrypted codes into a WebP image file. The WebP image is featured by being an independent, web friendly, and highly compressed file. In order to make the file describing its contents, we reserved some pixels as coded descriptive pixels. By this way, the now packed data exhibits its contents and type during image preview.

It is proven that the Data-In-Image format, regardless of being encrypted, occupies the least amount of storage space among other image formats that can be easily handled, stored, and shared through clouds and devices safely with a lower cost. For seismic data, the size of the WebP image comprises ~20% of the corresponding binary size with a bit-rate of ~5.6 b/s which is smaller than that of the Steim form, 27% and 8.9 b/s, respectively. Regarding the compression speed, it is found that the code compresses data with a rate of ~11,118 samples/s or ~ 44 Kbytes/s in average.

In addition, the data image is able to be digitally scanned and with some modifications can be remotely accessed like the quick response code, the thing that is not possible in the binary form. Moreover, the descriptive pixels in the image allow the implementations of smart tools to archive and classify data by machine learning and recognition algorithms.