

Tema Amplio de Investigación

Alexander Enrique Urieles Nieto
aeurieleesn@unal.edu.co

Maestría en Sistemas y Computación
Departamento de Ingeniería de Sistemas e Industrial
Universidad Nacional de Colombia

Tema Amplio de Investigación

De acuerdo con mi Interés de Investigación [1], el área de Compresión de Información es mi área de interés general de investigación, el cual hace parte del estudio de las Ciencias de la Computación y de la Teoría de la Información.

A lo largo de los años múltiples algoritmos han aparecido mejorando la calidad de la compresión especialmente LZ77 [2] y LZ78 [3], así como sus variaciones.

Sin embargo, junto con el aumento en la calidad de la compresión, la cantidad de información para almacenar no únicamente a nivel local sino a nivel mundial ha aumentado también, especialmente para las bases de datos. El progreso en las técnicas de compresión ha ayudado a aminorar la cantidad de espacio en disco que ocupa esta información.

Desgraciadamente para poder realizar búsquedas en esta información comprimida se necesitaba primero descomprimir y utilizar técnicas de búsquedas de subcadenas (como Sunday [4]) para encontrar las ocurrencias del patrón buscado en el texto comprimido.

Desde el trabajo presentado en [5] se inició una nueva rama de investigación denominada *Compressed Matching Problem*, que busca resolver el problema de búsqueda de ocurrencias de un patrón ya sea exacto, aproximado o complejos directamente en el texto comprimido sin necesidad de descompresión.

En *Compressed Matching Problem* hay una necesidad de encontrar mejores soluciones para los problemas de ocurrencias exactas, aproximadas y complejas que sean por lo menos igualmente de rápidas que en el problema de encontrar las ocurrencias en el texto descomprimido original.

Referencias

- [1] A. Urieles. Interés de Investigación. *Seminario de Investigación I*, 2011.
- [2] J. Ziv and A. Lempel. A universal algorithm for sequential data compression. *Information Theory, IEEE Transactions on*, 2003(3):337—343, 2002.

- [3] J. Ziv and A. Lempel. Compression of individual sequences via variable-rate coding. *Information Theory, IEEE Transactions on*, 24(5):530—536, 2003.
- [4] D.M. Sunday. A very fast substring search algorithm. *Communications of the ACM*, 33(8):132–142, 1990.
- [5] A. Amir and C. Benson. Efficient two-dimensional compressed matching. *Data Compression Conference, 1992.*, 2002.