ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

Design and Implementation of Data Compression with Hybrid to Optimized Size

Rajendra Choudhary, Dr. Aakash Saxena

M. Tech Scholar CS, Comucom Institute of Technology and Management, Jaipur, Rajasthan, India

Associate Professor, Comucom Institute of Technology and Management, Jaipur, Rajasthan, India

ABSTRACT: In the present era of communication over the internet, most of our data are transmitted through the HTTP requests. It is convenient for small files to send over the internet but the large data file needs to be compressed, because the present era is already suffering from the inadequate bandwidth. Therefore compressing the data with the highest density gives a significant change in the transmission of data but without ignoring the quality as the medical and various images need the lossless images but without losing the data. We have investigated one of the most important problems in multimedia applications is the storage and transmission of image, video and audio data. This made the field of developing image compression methods necessary and vital. Different image compression techniques were proposed to achieve high compression ratios and high image qualities in low computation time.

I. INTRODUCTION

The Web Compression on Wired and Wireless Networks, Overview of various data compression Algorithms, Problem Statement, Motivation, Objective and the scope of the work.

Overview of Web Compression on Network

Compression is the depletion in data size by converting it to a format that requires fewer bits. Most often compression is used to prune the storage space (on a hard drive, for example) or for shrinking the data transmitted over a network. By shrinking the size of data transferred, more bandwidth is available and transmission hits are reduced. Web-Compression is a foremost technique that is used to enhance the Web retrieval (hits) latency. This ability to compress the data on-the-fly empowers the network to serve entities to clients with the least amount of server-side bandwidth. The impact of this type of compression is that the amount of transmitted bytes is condensed and thus obtains a higher performance. Compression technology uses the various algorithms to remove extraneous/ repetitive information. Plethora of algorithms is defined for reducing the stream of data. Some are good in achieving compression without any loss and else are better at lossy compression. Since in web page losses in contents are not permissible, thus it is necessary to stick to the scope of lossless compression techniques only.

Example of the lossless compression algorithms:

- LZ77
- LZSS
- Arithmetic Coding
- Huffman Coding

Example of the lossless compression algorithms:

- Transform Coding
- Quantization
- Wavelet-Based Coding

After practicing the compression on the data file, the original information is rendered into a more compact and efficient format. This “compressed” payload can then be sent over the network. Besides the compressed data is destined to its destination, it can be uncompressed or reconstruct into original using extraction algorithms.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

Overview of Web Compression on Wireless Network

In recent times, the wireless networks have become very popular. Wireless LANs are being deployed on almost every area such as airports, conferences, etc. People have started using portable laptops to access the Internet and other resources using wireless networks while moving. Another area which has generated a lot of interest recently, is wireless ad-hoc networks.

Techniques for saving bandwidth are very significant in bandwidth constrained wireless networks. One of the key concepts that save bandwidth is compression of data packets. While it has been employed over low bandwidth point-to-point wired or wireless links, multiuser communication networks such as WLANs and WMNs (wireless mesh networks) do not use packet compression. Some dynamic data compressions are implemented to achieve compression efficiency and throughput.

With development in technologies microelectronics and wireless communication technology hardware manufacturing ability is also improved. Low-Power, inexpensive sensor nodes may be incorporated with information gathering, data dispensation and wireless communication functions [1]. The sensor nodes are usually power-driven by battery and deploy in harsh substantial environment. It is usually unfeasible to change the batteries and nodes, so the goal of data compression in wireless sensor network (WSN) is to condense the energy utilization and extend the network latencies.

II. PROPOSED SYSTEM STRUCTURE

This chapter describes in detail the proposed System's design constraints, architecture, compression stage and implementation details.

System Overview

The key purpose of proposed system is to provide users with efficient and fast Web information transmission in less time mostly where the same static content transmitted many times. The system includes a backward reference model and a entropy encoder. As an input it takes the data content from the Web server and as an output user receives the encoded Web pages which is again decoded by the Web browsers to retrieve the web pages.

The mechanism initially selects backward references that have shorter backward distances and then make them passes through the greedy parser over a window slider. After collecting all the references they pass through the entropy encoder using Huffman coding and output a compressed stream which would be transmitted over the network. An overall view of the proposed system is shown in Fig

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

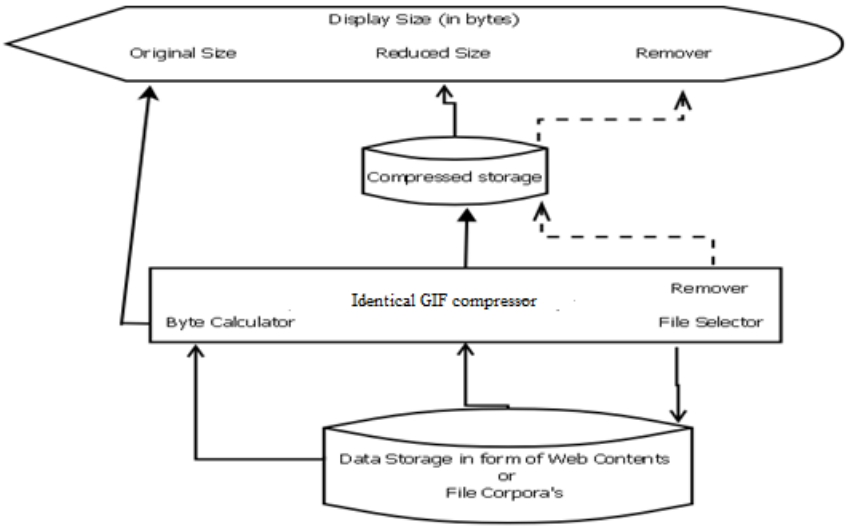


Fig. 1. System Overview of Identical GIF compressor

Proposed System Architecture

The system is a three-tier model.

- The first tier involves collecting the images required for information to be compressed for the transmissions and set the iteration counter to minimize the entropy.
- The second uses a steering system (model) to compare the references and route the matched statistics to analyzer for the scoring purposes.
- The third tier uses the above resulted scores to encode the data stream with the greedy heuristic and after passing through adaptive pallete, it yields compressed stream with best.

An overall architecture of the proposed system is shown in Fig. 3.2 on the next page.

- First box i.e. model show the comparison and selection of reference.
- Next the Encoder box shows the parsing operations.
- The Greedy Parsing and taking the shortest scans process as shown in Fig.2, but this process contains several steps which are explained in the later sections.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

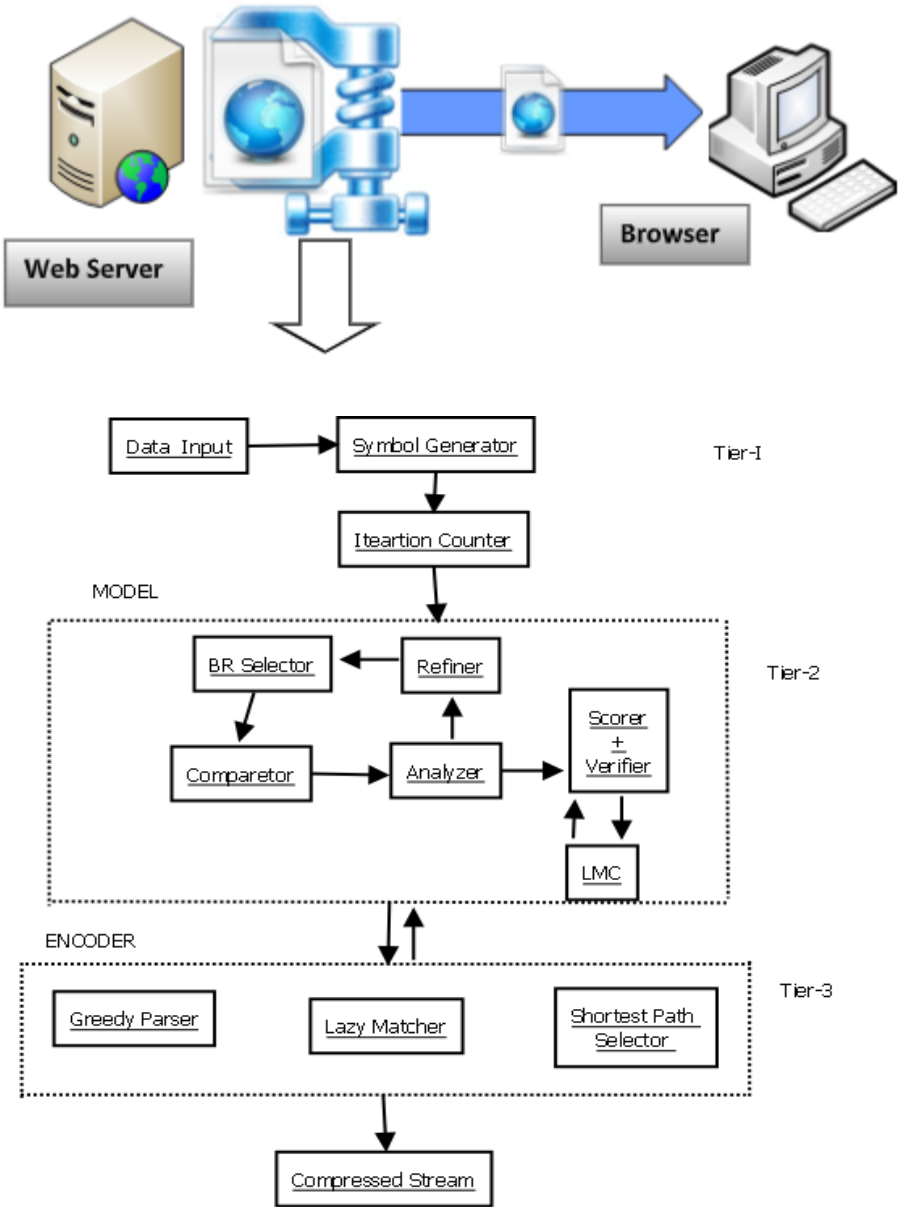


Fig. 2. System Architecture of the Proposed System

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

Proposed Identical GIF Algorithm

Algorithm 1. Algorithm for Proposed IdenticalGIFCompressor

```
finalize_optimizer(Gif_Stream *gfs, int optimize_flags)
{
 int i;

 if (background == TRANSP)
  gfs->background = (uint8_t)gfs->images[0]->transparent;

 /* 11.Mar.2010 - remove entirely transparent frames. */
 for (i = 1; i < gfs->nimages && !(optimize_flags & GT_OPT_KEEPEMPTY); ++i) {
  Gif_Image *gfi = gfs->images[i];
  if (gfi->width == 1 && gfi->height == 1 && gfi->transparent >= 0
      && !gfi->identifier && !gfi->comment
      && (gfi->disposal == GIF_DISPOSAL_ASIS
          || gfi->disposal == GIF_DISPOSAL_NONE
          || gfi->disposal == GIF_DISPOSAL_PREVIOUS)
      && gfi->delay && gfs->images[i-1]->delay) {
    Gif_UncompressImage(gfs, gfi);
    if (gfi->img[0][0] == gfi->transparent
        && (gfs->images[i-1]->disposal == GIF_DISPOSAL_ASIS
            || gfs->images[i-1]->disposal == GIF_DISPOSAL_NONE)) {
      gfs->images[i-1]->delay += gfi->delay;
      Gif_DeleteImage(gfi);
      memmove(&gfs->images[i], &gfs->images[i+1], sizeof(Gif_Image *) * (gfs->nimages - i - 1));
      --gfs->nimages;
      --i;
    }

 }
```

III. EXPERIMENTS

Two types of experiments are illustrated for the proposed system. These are: (i) the analytical experiment which gives the theoretical concept of the compression details, (ii) the empirical experiment describes the actual functioning of the compressions and their outputs. Compressing operations try to save as little information as possible which represent the compete message, attempting to retain high precision and relevancy of the results.

3.1 Analytical Experiment

An important benefit of the adaptive process is the ability to automatically obtain images symbols for matching Web server. This allows the system to identify longest matches without having to maintain a massive dictionary or database of symbols covering a wide range of categories. In this project. If the query gets more specific, then a smaller number of relevant content can be ingressed to compress. Based on the assumption mentioned above, one of the first and most important steps in the process is to move for matching current symbols with the previous one.

For detailed analysis, let's take the input the stream as "abcdefghi" at very start so it has not found any backward references. Now when the next line or the stream is "abcdegghi" then it evaporates the distance and length as -10 where '-'shows the back movement. As with the same when meet with the next stream as "00000abcde" then it has multiple choice for the references to encode zeros – (1) then distance and length are -1 and 4 (2) and -10 and 5 respectively. Next when "000cdefgi" comes into track then to encode '000' it gets back matched with distance -8 and length 3 ;

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

distance -9 and length 3 and distance -10 and length 3 each with different possible tuples. So for this the large distances are being selected which is more probable statistically and leads to smaller entropy.
As the iterations are subsequently searches the matches, they get analyzed and store their only references and some are not compared that would be handled by lazy scans so that dense compressions can be achieved. Now when the iterations being finished, the scores will be indexed or ranked and best matches are taken first to encode for speed up process and after that other can be encoded. After all these processes a dense and highest compressed output will be emerged with the format that is compatible with all the other decompresses as well as compatible with the decoders of the browsers. The complete process summarizes into a flowchart as shown in Fig.

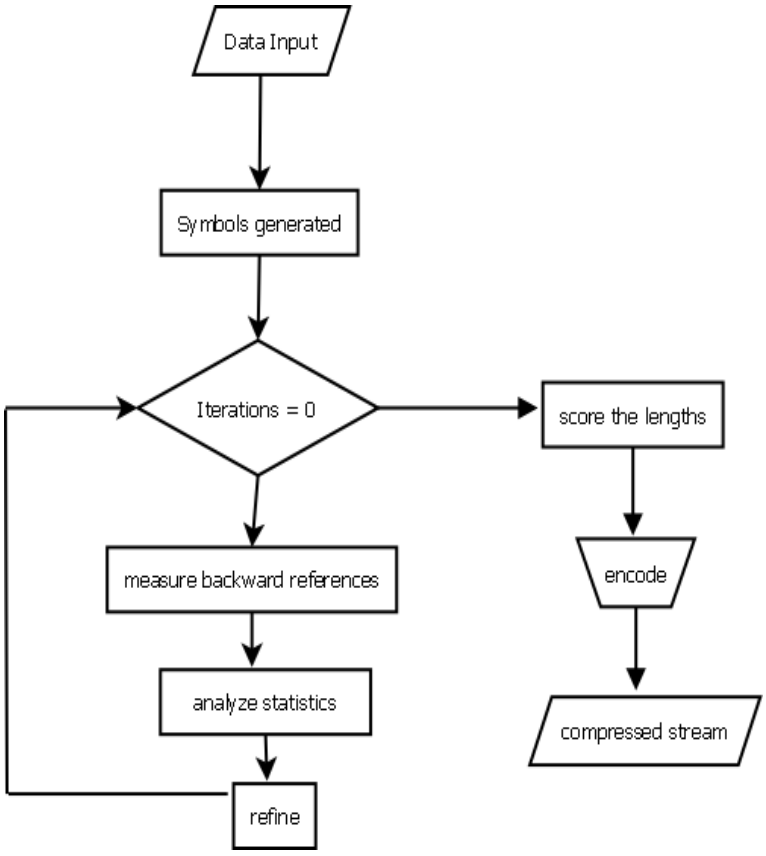


Fig. 3. Steps involved in the Proposed Architecture

IV. EMPIRICAL EXPERIMENT

How the system interface compresses the file or the web content files and at where the storage resides. In this chapter the different modules or the sub models are shown that helps in compressing the Web contents. Here demo structure of the whole idea of the proposed compression technique has been built with the detailed model submodules viz. various compressor and the bytes counter with the content remover.
There are gif image benchmarks that have been taken for the purpose of pointing out the performance of a compression algorithm. These files exist because it wouldn't be reasonable neither to examine performance on a few image chosen by programmers nor on a set of image of the same type. Every algorithm can measure its capability on a corpus, and in [40], it is possible to find the results of years and years of experimentation.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

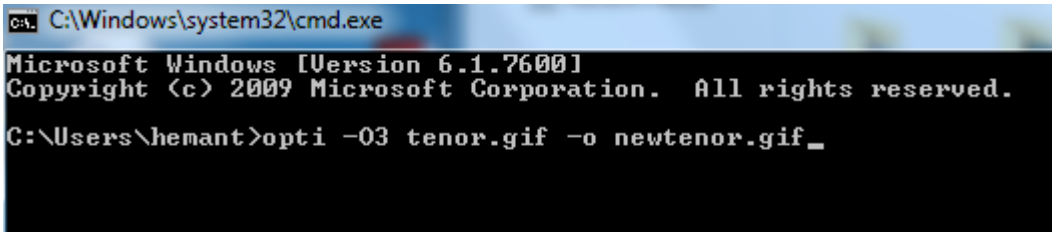
Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

In particular just focus on different images corpus (see [40] and [41]) and of GIF set, a newly used tenor benchmark (see [42] and [43]). However, the results will be also given on other imgaes that, according to it, represent a good benchmark for testing a compression algorithm, even in its extreme cases. Each of the compressors has its own hold storage where it compresses the GIF which needs to be transmitting over the networks which are discussed as follows:

IdenticalGIF

It is based on the proposed technique which is an improvement over most famous compressors like LZW etc. For experimenting with it a separate storage has made for the compressing with the IdenticalGIF where the entire corpus sets are being compressed or say the images which needs to be compressed are being prepared here for transmission. After transmitting all the files over the network it clears the storage by cleaner or remover. So that the storage cannot saved the repeated copies of the same set in different formats. Otherwise it will again create space problems on server.
As starting with the proposed system provides a simple images ingressing command interface where it shows the modules with various technologies, from where very newly proposed compressor identicalGIF is used.



V. CONCLUSION

The industry of web searching continues to grow along with the rapidly growing World Wide Web. The proposed Idetnical GIF based on adaptive and greedy parsing Algorithm makes the compressing easier by providing some sort of guidance for inexperienced users and for users of the Internet in general. This project has achieved two goals:

- ✓ transmitting the big contents over Internet in very less time
- ✓ reducing the size with same quality lossless .

Compressing the data pages or the image web content allows the user to retrieve more relevant information in fastest way, because adding more files or the word contents to the original ones narrows down the compressed results, making results outperforms and more efficient, and applying this new IdenticalGIF compression process to a Web service makes the user to reach to the appropriate result in a less number of hits.
Experimental results show that the proposed system suggest the phenomenon with 26% - 35% compression ratio which is much significant for the content over HTTP because the bandwidth costs much for the increasing time.

REFERENCES

1. Institute of Electrical Electronics Engineers, “Ten emerging technologies that will change your world”, *IEEE Engineering Management Review*, vol. - 32, pp. 20–30, 2004.
2. Stephan Rein and Martin Reisslein., “Performance evaluation of the fractional wavelet filter: A low-memory image wavelet transform for multimedia sensor networks”, *Ad Hoc Networks, vol.* - 9, no.-4, pp. 482-496, June 2011. DOI=10.1016/j.adhoc.2010.08.004, http://dx.doi.org/10.1016/j.adhoc.2010.08.004.
3. Z. J. Xie, Lei Wang, Y. H. Liu, and H. Chen, “Voronoi tessellation based multiscale data compression for sensor networks”, In *Proceedings of the 5th WSEAS international conference on Telecommunications and informatics* (TELE-INFO'06), Metin Demiralp, Aydin Akan, and Nikos Mastorakis (Eds.). World Scientific and Engineering Academy and Society (WSEAS), Stevens Point, Wisconsin, USA, pp. 476-486, 2006.
4. Reinhardt, A., Christin, D., Hollick, M., Schmitt, J., Mogre, P.S., Steinmetz, R., “Trimming the Tree: Tailoring Adaptive Huffman Coding to Wireless Sensor Networks”, In Silva, J.S., Krishnamachari, B., Boavida, F. (eds.) EWSN 2010. LNCS, vol. - 5970, pp. 33-48, Springer, Heidelberg, 2010.
5. Yeo, M., Seong, D., Cho, Y., Yoo, J., “Huffman Coding Algorithm for Compression of Sensor Data in Wireless Sensor Networks”, *In ICHIT 2009 - International Conference on Convergence and Hybrid Information Technology*, pp. 296-301, 2009.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

6. Shangfeng Mo, Hong Chen, Yinglong, Li, "Topology-Based Data Compression in Wireless Sensor Networks", *13th International Conference*, WAIM 2012, Harbin, China, vol.-7418, pp. 22-34, August 18-20, 2012.
7. Zhigang Liu; Saifullah, Y.; Greis, M.; Sreemanthula, S., "HTTP compression techniques", *Wireless Communications and Networking Conference*, IEEE, vol.-4, pp. 2495-2500, 13-17 March 2005. doi: 10.1109/WCNC.2005.1424906
8. Mark Nelson and Jean-loup Gailly, "The Data Compression Book", *ISBN 1558514341*, Wiley, 1995.