x3: Lossless Data Compressor

David Barina and Ondrej Klima

Brno University of Technology, Brno, Czech Republic
{ibarina,iklima}@fit.vutbr.cz

x3[1] is a lossless optimizing dictionary-based data compressor. The algorithm uses a combination of a dictionary, context modeling, and arithmetic coding. Optimization adds the ability to find the most appropriate parameters for each file. Even without optimization, x3 can compress data with a compression ratio comparable to the best dictionary compression methods like LZMA, zstd, or Brotli.

The main goal in x3 design was to achieve maximal compression ratio, even at the cost of large memory requirements and long compression times. The core of the algorithm is a dictionary. Unlike methods based on LZ77, the x3 algorithm looks for suitable fragments in a search window, which it then explicitly inserts into the dictionary. Each fragment stored in the dictionary gets a unique number (index), and a sequence of these indexes then forms the compressed stream. The backend of the x3 algorithm consists of a context arithmetic encoder. The x3 offers the possibility of optimization (search for suitable parameters for each file). This optimization only affects the compression process and does not involve the need to transmit additional data to the decoding side. The fundamental parameters include the size of the search window and the maximum number of matches in this window. Below, we further compare the x3 algorithm with other dictionary methods used in practice. The well-known Silesia corpus is used for this comparison (best results in bold). Experimental evaluation shows that x3 can compress files with a higher compression ratio than all other state-of-the-art dictionary methods (including Brotli, zstd, and LZMA).

File	LZ4	gzip	xz	zstd	Brotli	x3
dickens	2.2948	2.6461	3.6000	3.5765	3.6044	**3.7168**
mozilla	2.3176	2.6966	**3.8292**	3.3769	3.6922	2.7432
mr	2.3472	2.7138	3.6231	3.2132	3.5317	**4.0364**
nci	9.1071	11.2311	**23.1519**	20.7925	22.0780	19.1103
ooffice	1.7349	1.9907	**2.5346**	2.3587	2.4818	2.0668
osdb	2.5290	2.7138	3.5456	3.2855	3.5812	**3.6151**
reymont	3.1345	3.6396	5.0374	4.9060	4.9747	**5.1010**
samba	3.5122	3.9950	**5.7778**	5.5267	5.7367	4.1871
sao	1.2639	1.3613	**1.6386**	1.4479	1.5812	1.5042
webster	2.9554	3.4372	4.9540	4.8970	4.9188	**4.9685**
xml	6.9277	8.0709	12.2910	11.8004	**12.4145**	9.2249
x-ray	1.1798	1.4035	1.8868	1.6457	1.8096	**1.9649**

---

[1]https://github.com/xbarin02/x3-compressor