

Universidade Federal da Paraíba
Centro de Informática

Disciplina: Introdução à Teoria da Informação

Equipe:

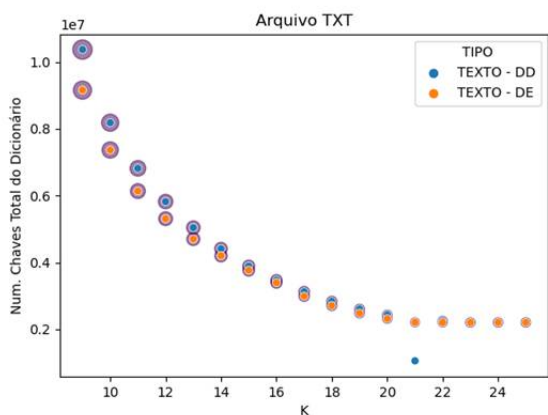
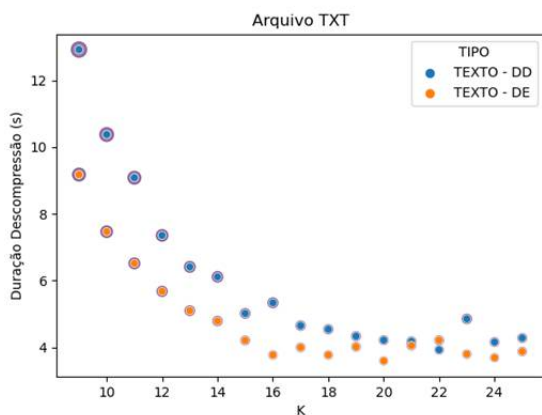
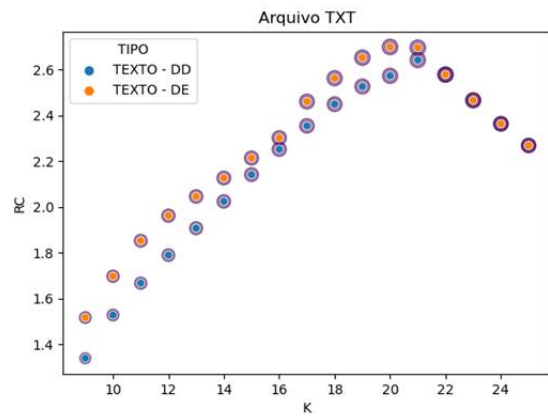
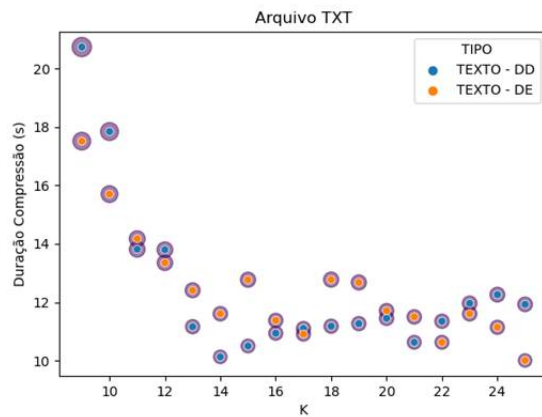
Everaldo da Costa Mendes	0160112738
José Ítalo Alves de Oliveira Vitorino	2016005182
Sérgio Ricardo Araújo do Nascimento	20200022201

O objetivo do presente relatório é apresentar os resultados obtidos através de um compactador LZW desenvolvido na linguagem Python pela equipe acima relacionada para analisar dois arquivos um de cerca de 16 MB no formato texto e outro arquivo de vídeo tipo MP4.

Implementamos duas opções do LZW, uma com o dicionário estático, ou seja, quando a quantidade de itens no dicionário for preenchido, o dicionário não será mais atualizado. A outra implementação foi realizada com o dicionário dinâmico, ou seja, quando o dicionário encher, ele será esvaziado, reinicializado (com os 256 caracteres do alfabeto – caracteres 0 a 255) e então passará a ser atualizado novamente.

Fizemos análise das compactações (tamanhos e tempos) nos dos arquivos e em ambas as implementações do LZW. O K utilizado variou de 9 a 25.

Os resultados seguem abaixo:

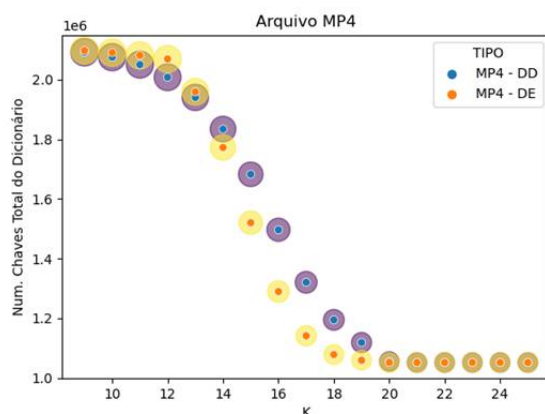
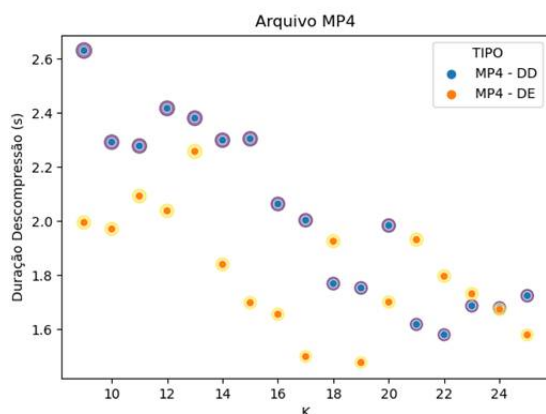
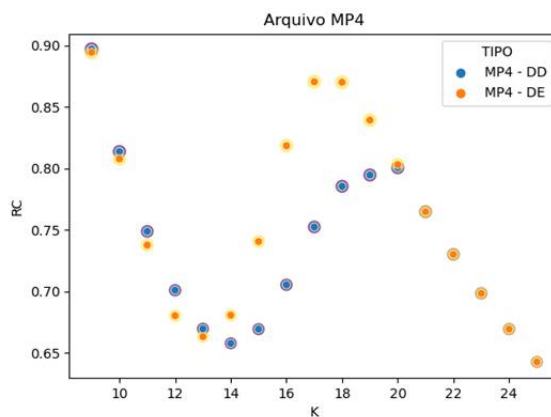
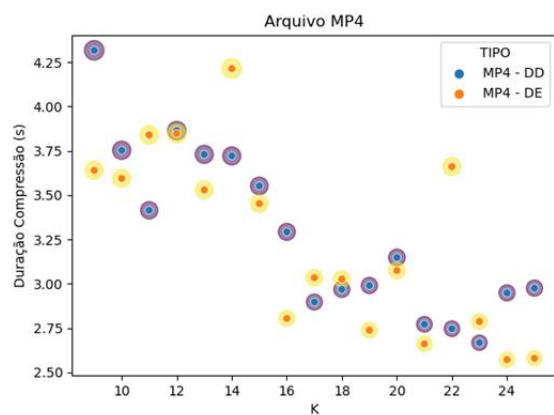


Nos gráficos acima podemos observar o seguinte:

K x Duração de Compressão e K x Duração de Descompressão: À medida que o K aumenta a duração da compressão também aumenta, isto se dá por que a quantidade de chaves do dicionário aumenta e com isto o tamanho das chaves também aumenta.

K x RC: À medida que o K aumenta a RC também aumenta devido ao tamanho das palavras do dicionário, mas em um certo ponto (neste caso K = 22) a RC começa a diminuir, isto por que os tamanhos das chaves aumentam e faz com que ocupe mais espaço.

K x Núm. Chaves Total do Dicionário: Quando o K aumenta a quantidade total de chaves utilizadas diminui por que há cada vez mais chaves no dicionário a serem utilizadas.



Nos gráficos acima podemos observar o seguinte:

K x Duração de Compressão e K x Duração de Descompressão: Aqui não há um padrão da relação, isto porque trata-se de arquivo MP4 que já é compactado.

K x RC: Por se tratar de um arquivo de entrada já compactado, o compactador não conseguiu um RC maior que um. O gráfico mostra que o RC começa diminuindo, depois aumenta e termina diminuindo, isto enquanto o K aumenta.

K x Núm. Chaves Total do Dicionário: Quando o K aumenta a quantidade total de chaves utilizadas diminui por que há cada vez mais chaves no dicionário a serem utilizadas.

Arquivo CORPUS16MB.TXT com Dicionário Dinâmico

K	Tamanho Inicial	Tamanho Final	RC	Incrmento do RC	Melhorou Compactação ?	Num. Chaves do Último Dicionário	Num. Chaves Total do Dicionário	Duração Compressão (s)	Duração Descompressão (s)
9	15.637.070	11.668.029	1,3402			508	10.371.581	20,739	12,917
10	15.637.070	10.228.573	1,5288	1,1407	Sim	841	8.182.858	17,845	10,373
11	15.637.070	9.374.704	1,6680	1,0911	Sim	1.453	6.817.966	13,812	9,081
12	15.637.070	8.733.656	1,7904	1,0734	Sim	1.252	5.822.437	13,797	7,355
13	15.637.070	8.195.516	1,9080	1,0657	Sim	4.289	5.043.394	11,161	6,412
14	15.637.070	7.721.783	2,0251	1,0614	Sim	9.758	4.412.447	10,128	6,115
15	15.637.070	7.299.589	2,1422	1,0578	Sim	24.441	3.893.114	10,500	5,020
16	15.637.070	6.940.646	2,2530	1,0517	Sim	10.738	3.470.323	10,936	5,338
17	15.637.070	6.639.231	2,3553	1,0454	Sim	115.831	3.124.344	11,098	4,654
18	15.637.070	6.382.960	2,4498	1,0401	Sim	218.246	2.836.871	11,177	4,546
19	15.637.070	6.187.721	2,5271	1,0316	Sim	509.483	2.605.356	11,266	4,338
20	15.637.070	6.075.640	2,5737	1,0184	Sim	333.871	2.430.256	11,448	4,220
21	15.637.070	5.916.134	2,6431	1,0270	Sim	6.669	1.054.734	10,627	4,186
22	15.637.070	6.063.041	2,5791	0,9758	Não	157.124	2.253.765	11,347	3,930
23	15.637.070	6.338.634	2,4669	0,9565	Não	2.204.997	2.204.742	11,961	4,855
24	15.637.070	6.614.226	2,3642	0,9583	Não	2.204.997	2.204.742	12,258	4,163
25	15.637.070	6.889.819	2,2696	0,9600	Não	2.204.997	2.204.742	11,926	4,284
MÉDIA		7.486.464,82	2,1697	1,0347		471.147	4.054.921,88	12,472	5,987
MEDIANA		6.889.819,00	2,2696	1,0428		24.441	3.124.344,00	11,347	4,855
DESVIO PADRÃO		3.378.704,69	0,6572	0,1278		1.558.809	5.774.827,24	6,232	6,104
MAIOR		11.668.029	2,6431	1,1407		2.204.997	10.371.581	20,739	12,917
MENOR		5.916.134	1,3402	0,9565		508	1.054.734	10,128	3,930

Arquivo CORPUS16MB.TXT com Dicionário Estático

K	Tamanho Inicial	Tamanho Final	RC	Incrmento do RC	Melhorou Compactação ?	Num. Chaves do Último Dicionário	Num. Chaves Total do Dicionário	Duração Compressão (s)	Duração Descompressão (s)
9	15.637.070	10.302.668	1,5178			512	9.157.927	17,517	9,175
10	15.637.070	9.208.047	1,6982	1,1189	Sim	1.024	7.366.437	15,703	7,466
11	15.637.070	8.437.402	1,8533	1,0913	Sim	2.048	6.136.292	14,173	6,517
12	15.637.070	7.966.400	1,9629	1,0591	Sim	4.096	5.310.933	13,349	5,679
13	15.637.070	7.639.697	2,0468	1,0428	Sim	8.192	4.701.352	12,408	5,101
14	15.637.070	7.350.504	2,1273	1,0393	Sim	16.384	4.200.288	11,604	4,791
15	15.637.070	7.060.950	2,2146	1,0410	Sim	32.768	3.765.840	12,769	4,211
16	15.637.070	6.790.094	2,3029	1,0399	Sim	65.536	3.395.047	11,376	3,776
17	15.637.070	6.353.113	2,4613	1,0688	Sim	131.072	2.989.700	10,907	4,002
18	15.637.070	6.100.279	2,5633	1,0414	Sim	262.144	2.711.235	12,774	3,776
19	15.637.070	5.893.869	2,6531	1,0350	Sim	524.288	2.481.629	12,676	4,021
20	15.637.070	5.791.765	2,6999	1,0176	Sim	1.048.576	2.316.706	11,707	3,599
21	15.637.070	5.798.528	2,6967	0,9988	Não	2.097.152	2.208.963	11,496	4,061
22	15.637.070	6.063.041	2,5791	0,9564	Não	2.204.997	2.204.742	10,627	4,215
23	15.637.070	6.338.634	2,4669	0,9565	Não	2.204.997	2.204.742	11,603	3,801
24	15.637.070	6.614.226	2,3642	0,9583	Não	2.204.997	2.204.742	11,139	3,694
25	15.637.070	6.889.819	2,2696	0,9600	Não	2.204.997	2.204.742	10,005	3,880
MÉDIA		7.094.060,94	2,2634	1,0266		765.516	3.856.548,06	12,461	4,810
MEDIANA		6.790.094,00	2,3029	1,0396		131.072	2.989.700,00	11,707	4,061
DESVIO PADRÃO		2.413.248,67	0,5316	0,1123		1.558.806	4.916.644,26	5,312	3,744
MAIOR		10.302.668	2,6999	1,1189		2.204.997	9.157.927	17,517	9,175
MENOR		5.791.765	1,5178	0,9564		512	2.204.742	10,005	3,599

Arquivo DISCO RIGIDO.MP4 com Dicionário Dinâmico

K	Tamanho Inicial	Tamanho Final	RC	Incrmento do RC	Melhorou Compactação ?	Num. Chaves do Último Dicionário	Num. Chaves Total do Dicionário	Duração Compressão (s)	Duração Descompressão (s)
9	2.111.047	2.353.462	0,8970			445	2.091.966	4,317	2,631
10	2.111.047	2.594.499	0,8137	0,9071	Não	718	2.075.599	3,753	2,293
11	2.111.047	2.819.518	0,7487	0,9202	Não	765	2.050.558	3,414	2,279
12	2.111.047	3.012.224	0,7008	0,9360	Não	3.924	2.008.149	3,864	2,418
13	2.111.047	3.153.376	0,6695	0,9552	Não	4.410	1.940.539	3,730	2,381
14	2.111.047	3.210.143	0,6576	0,9823	Não	12.158	1.834.367	3,721	2,300
15	2.111.047	3.154.924	0,6691	1,0175	Sim	24.769	1.682.626	3,552	2,305
16	2.111.047	2.993.372	0,7052	1,0540	Sim	60.781	1.496.686	3,293	2,064
17	2.111.047	2.806.008	0,7523	1,0668	Sim	12.569	1.320.474	2,897	2,004
18	2.111.047	2.687.756	0,7854	1,0440	Sim	147.261	1.194.558	2,968	1,770
19	2.111.047	2.656.635	0,7946	1,0117	Sim	70.774	1.118.583	2,990	1,754
20	2.111.047	2.636.835	0,8006	1,0075	Sim	6.669	1.054.734	3,148	1,984
21	2.111.047	2.760.366	0,7648	0,9552	Não	1.051.823	1.051.568	2,771	1,619
22	2.111.047	2.891.812	0,7300	0,9545	Não	1.051.823	1.051.568	2,746	1,581
23	2.111.047	3.023.258	0,6983	0,9565	Não	1.051.823	1.051.568	2,667	1,687
24	2.111.047	3.154.704	0,6692	0,9583	Não	1.051.823	1.051.568	2,948	1,680
25	2.111.047	3.286.150	0,6424	0,9600	Não	1.051.823	1.051.568	2,975	1,725
MÉDIA		2.893.826,00	0,7353	0,9804		329.668	1.478.039,94	3,280	2,028
MEDIANA		2.891.812,00	0,7300	0,9592		24.769	1.320.474,00	3,148	2,004
DESVIO PADRÃO		659.510,01	0,1800	0,0374		743.437	735.672,48	0,949	0,641
MAIOR		3.286.150	0,8970	1,0668		1.051.823	2.091.966	4,317	2,631
MENOR		2.353.462	0,6424	0,9071		445	1.051.568	2,667	1,581

Arquivo DISCO RIGIDO.MP4 com Dicionário Estático

K	Tamanho Inicial	Tamanho Final	RC	Incrmento do RC	Melhorou Compactação ?	Num. Chaves do Último Dicionário	Num. Chaves Total do Dicionário	Duração Compressão (s)	Duração Descompressão (s)
9	2.111.047	2.359.912	0,8945			512	2.097.699	3,640	1,996
10	2.111.047	2.613.378	0,8078	0,9030	Não	1.024	2.090.702	3,594	1,972
11	2.111.047	2.862.104	0,7376	0,9131	Não	2.048	2.081.530	3,840	2,093
12	2.111.047	3.104.067	0,6801	0,9220	Não	4.096	2.069.378	3,848	2,039
13	2.111.047	3.183.843	0,6630	0,9749	Não	8.192	1.959.288	3,530	2,258
14	2.111.047	3.102.087	0,6805	1,0264	Sim	16.384	1.772.621	4,214	1,841
15	2.111.047	2.850.722	0,7405	1,0882	Sim	32.768	1.520.385	3,452	1,699
16	2.111.047	2.579.156	0,8185	1,1053	Sim	65.536	1.289.578	2,804	1,656
17	2.111.047	2.424.810	0,8706	1,0637	Sim	131.072	1.141.087	3,034	1,500
18	2.111.047	2.426.148	0,8701	0,9994	Não	262.144	1.078.288	3,026	1,926
19	2.111.047	2.515.244	0,8393	0,9646	Não	524.288	1.059.050	2,738	1,478
20	2.111.047	2.628.930	0,8030	0,9568	Não	1.048.576	1.051.572	3,075	1,701
21	2.111.047	2.760.366	0,7648	0,9524	Não	1.051.823	1.051.568	2,661	1,932
22	2.111.047	2.891.812	0,7300	0,9545	Não	1.051.823	1.051.568	3,661	1,797
23	2.111.047	3.023.258	0,6983	0,9565	Não	1.051.823	1.051.568	2,787	1,732
24	2.111.047	3.154.704	0,6692	0,9583	Não	1.051.823	1.051.568	2,571	1,674
25	2.111.047	3.286.150	0,6424	0,9600	Não	1.051.823	1.051.568	2,579	1,580
MÉDIA		2.809.805,35	0,7594	0,9812		432.691	1.439.354,00	3,238	1,816
MEDIANA		2.850.722,00	0,7405	0,9592		131.072	1.141.087,00	3,075	1,797
DESVIO PADRÃO		654.949,17	0,1783	0,0403		743.389	739.726,32	0,750	0,294
MAIOR		3.286.150	0,8945	1,1053		1.051.823	2.097.699	4,214	2,258
MENOR		2.359.912	0,6424	0,9030		512	1.051.568	2,571	1,478