**Clustering Results on 16S Environmental sets using Threshold #1**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Algorithm | Metric | | 53R | 55R | 112R | 115R | 137 | 138 |
| CDMLz4 | # Clu | | 320 | 288 | 507 | 324 | 306 | 308 |
| W.Sim | | 94.30 | 93.35 | 92.01 | 94.30 | 93.81 | 93.91 |
| Time(s) | | 6.9 | 5.1 | 5.5 | 5.3 | 9.2 | 6.4 |
| CDMSequitur | # Clu | 346 | | 292 | 528 | 322 | 306 | 319 |
| W.Sim | 93.93 | | 93.07 | 92.41 | 92.42 | 93.79 | 93.99 |
| Time(s) | 59.4 | | 44.9 | 59.5 | 333.0 | 71.8 | 58.3 |
| Lz4 | # Clu | 517 | | 448 | 775 | 509 | 456 | 465 |
| W.Sim | 96.63 | | 95.39 | 96.00 | 95.73 | 96.08 | 95.91 |
| Time(s) | 21.5 | | 14.3 | 20.8 | 168.9 | 24.0 | 19.2 |
| MC-MinH | # Clu | 1165 | | 1077 | 1634 | 1156 | 1020 | 1042 |
| W.Sim | 96.90 | | 92.45 | 91.18 | 93.33 | 95.86 | 93.10 |
| Time(s) | 2.5 | | 2.1 | 3.3 | 3.0 | 2.7 | 2.5 |
| MC-LSH | #Clu | 1172 | | 1199 | 1795 | 1205 | 1041 | 1072 |
| W.Sim | 96.90 | | 93.12 | 91.33 | 93.50 | 95.86 | 93.10 |
| Time(s) | 161.0 | | 183.0 | 317.0 | 188.0 | 172.0 | 175.0 |
| UCLUST | #Clu | 1062 | | 992 | 1561 | 1071 | 900 | 923 |
| W.Sim | 96.67 | | 91.67 | 91.02 | 93.33 | 93.50 | 92.82 |
| Time(s) | 2.0 | | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 |
| CD-HIT | #Clu | 824 | | 716 | 1196 | 820 | 712 | 725 |
| W.Sim | 92.56 | | 90.80 | 90.61 | 93.33 | 91.82 | 90.16 |
| Time(s) | 3.6 | | 3.1 | 3.9 | 3.8 | 3.2 | 3.1 |
| ESPRIT | #Clu | 940 | | 859 | 1361 | 970 | 818 | 832 |
| W.Sim | 93.12 | | 91.35 | 90.88 | 93.33 | 91.82 | 90.16 |
| Time(s) | 283.0 | | 266.0 | 537.0 | 348.0 | 280.0 | 296.0 |
| DOTUR | #Clu | 1241 | | 1258 | 1854 | 1279 | 1096 | 1121 |
| W.Sim | 96.95 | | 94.06 | 91.33 | 93.50 | 95.86 | 93.10 |
| Time(s) | 5129.0 | | 3511.0 | 5567.0 | 9237.0 | 6563.0 | 5618.0 |
| Mothur | #Clu | 1238 | | 1256 | 1853 | 1278 | 1094 | 1119 |
| W.Sim | 96.95 | | 94.06 | 91.33 | 93.50 | 95.86 | 93.10 |
| Time(s) | 10130.0 | | 5940.0 | 12303.0 | 13501.0 | 12861.0 | 12310.0 |

Table   
CDMLz4 has the fastest runtime out of the three algorithms created in this project, and it falls second only to MC-MinH when compared to all used algorithms. The weighted similarity of the three algorithms closely matches the weighted similarity of the standard algorithms, as this threshold was chosen by weighted similarity matching. However the number of clusters were much lower than the expected values.