YagoVB

总带有坐标信息节点数：4774796

总带有时间和文本信息节点数：812532

|  |  |  |  |
| --- | --- | --- | --- |
|  | Temporal Inverted Index | 记录wid对应的<nid,date> index | 记录将原图转化为一个DAG图后，原图节点与DAG图的映射关系文件 |
| 时间 | 2s | 10s | 3m44s |
| 大小 | 22M | 34M | 137M |

|  |  |  |  |
| --- | --- | --- | --- |
| TF-label | rtree | 记录每个带有时间属性的节点，其最小和最大时间文件 | 记录每个pid对应的rtree node id，及每个rtreeNode对应的父节点id |
| 2m15s | 31m12s | 3s | 2m4s |
| 769M | 274M | 13M | 19M |

|  |  |  |  |
| --- | --- | --- | --- |
| 记录Rtree node到带有时间和文本属性节点的可达性文件 | 记录所有带坐标信息节点id（即pid） | 记录所有wid |  |
| 7m48s | 未记 | 未记 |  |
| 364M | 19M | 2M |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| alpha\_len文件记录   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | Len | Alpha=1 | | Alpha=2 | | Alpha=3 | | | 时间 | 大小 | 时间 | 大小 | 时间 | 大小 | | 不限制len | 未知 | 541M | 未知 | 5G | 未知 | 21G | | 100000 | 2m52s | 47M | 20m20s | 109M | 2h10m | 130M | | 1000000 | 3m5s | 140M | 23m6s | 856M | 2h38m | 1.8G | | 10000000 | 3m7s | 170M | 24m40s | 1.3G | 3h5m | 4.9G | |

WordFrequency（记录wid到pid的可达性及最短路径长度的index）

词频阈值 文件大小

0 38G

50 6.0G

100 3.9G

250 1.9G

500 929M

1000 448M

10000 32M

100000 3.2M

1000000 332k

DBpediaVB

总带有坐标信息节点数：883665

总带有时间和文本信息节点数：1138751

|  |  |  |  |
| --- | --- | --- | --- |
|  | Temporal Inverted Index | 记录wid对应的<nid,date> index | 记录将原图转化为一个DAG图后，原图节点与DAG图的映射关系文件 |
| 时间 | 26s | 2m5s | 未记 |
| 大小 | 118M | 292M | 128M |

|  |  |  |  |
| --- | --- | --- | --- |
| TF-label | rtree | 记录每个带有时间属性的节点，其最小和最大时间文件 | 记录每个pid对应的rtree node id，及每个rtreeNode对应的父节点id |
| 2m15s | 16m | 未记 | 未记 |
| 769M | 156M | 18M | 3.4M |

|  |  |  |  |
| --- | --- | --- | --- |
| 记录Rtree node到带有时间和文本属性节点的可达性文件 | 记录所有带坐标信息节点id（即pid） | 记录所有wid |  |
| 58m | 未记 | 未记 |  |
| 1.8G | 3.4M | 3.7M |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| alpha\_len文件记录   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | Len | Alpha=1 | | Alpha=2 | | Alpha=3 | | | 时间 | 大小 | 时间 | 大小 | 时间 | 大小 | | 不限制len | 未知 | 2.3G | 未知 | 8.5G | 未知 | 39G | | 100000 | 12m40s | 110M | 38m3s | 277M | 2h59m | 572M | | 1000000 | 13m10s | 270M | 40m9s | 914M | 3h4m | 2.5G | | 10000000 | 15m15s | 935M | 1h11m | 2.9G | 3h14m | 7.3G | |

WordFrequency（记录wid到pid的可达性及最短路径长度的index）

词频阈值 文件大小

100 55G

250 30G

500 19G

1000 12G

10000 1.7G

100000 53M

1000000 332K