1. degree分布

数据集中实际有100000个结点，其中67149个degree大于0，其余degree为0

最大的degree为12506

最大的100个degree值如下：

1 - 10 [12506. 12012. 11179. 7679. 6244. 5837. 5006. 4511. 2408.]

11 - 20 [2223. 1985. 1740. 1486. 1200. 1148. 1142. 1020. 957.]

21 - 30 [728. 632. 630. 626. 577. 576. 565. 564. 546.]

31 - 40 [488. 473. 431. 292. 272. 264. 244. 241. 231.]

41 - 50 [200. 191. 172. 151. 143. 141. 131. 131. 97.]

51 - 60 [95. 95. 91. 89. 86. 84. 79. 77. 76.]

61 - 70 [75. 71. 66. 64. 59. 58. 55. 54. 51.]

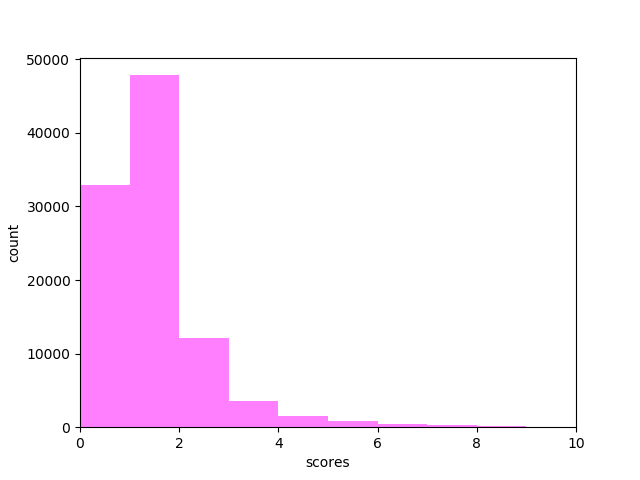
71 - 80 [50. 49. 48. 48. 47. 47. 45. 45. 45.]

81 - 90 [43. 40. 40. 39. 37. 34. 34. 33. 33.]

91 - 100 [32. 32. 32. 31. 31. 30. 30. 29. 29.]

其中第3、4和第9、10下降较多

绝大多数点的degree集中在5以内。



1. 通道

实验中选取了degree最大的m个结点形成通道，结果如下：

（实验结果已剔除degree为0的点，总结点数为67149）

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| m | 未划分进任何通道的结点数 | 与其他任何通道均不相连的通道数 | 通道的degree最小值 | 整张图的连通分支数 | BFS生成树的深度 |
| 4 | 30541 | 0 | 3 | 1 | 2 |
| 9 | 15386 | 0 | 8 | 1 | 2 |
| 100 | 1771 | 1 (76) | 0 | 2 | 3 |
| 500 | 1198 | 1 (290) | 0 | 2 | 5 |
| 1000 | 1035 | 1 (290) | 0 | 2 | 5 |
| 2000 | 827 | 1 (290) | 0 | 2 | 4 |
| 3357 | 616 | 1 (290) | 0 | 2 | 4 |

其中孤立的通道数量很少，而且基本都是同一个通道。

整张图的边非常多，随着m的上升，边的数量快速增加，目测是平方的关系，所以当m达到2000以后，BFS生成树的运行时间就很长。

1. 将孤立的点划分至最近的通道

|  |  |  |  |
| --- | --- | --- | --- |
| m | 划分失败的结点数 | 与其他任何通道均不相连的通道数 | 通道的degree最小值 |
| 4 | 0 | 0 | 3 |
| 9 | 0 | 0 | 8 |
| 100 | 0 | 1 (76) | 0 |
| 500 | 0 | 1 (290) | 0 |
| 1000 | 0 | 1 (290) | 0 |
| 2000 | 0 | 1 (290) | 0 |
| 3357 | 0 | 1 (290) | 0 |

可以看出，所有degree非零的结点都可以划进通道内，并且由于通道本身的连通性已经很好，所以在连通性上基本没有变化。