CSC 591, Network Properties in Spark GraphFrames

Student Id: 200261202

Mkhan8@ncsu.edu

**Solution 1:**

**1.**  4 random graphs were generated and tested for the power law for the corresponding degree distribution. The results are as follows:

|  |  |  |
| --- | --- | --- |
| **Graph** |  | **Scalefree** |
| Gnm1 | 2.88 | Yes |
| Gnm2 | 9.62 | No |
| Gnp2 | 54.58 | No |
| Gnp1 | 4.93 | No |

The degree distribution of the graphs are as follows:

1. Gnm1:

|  |  |
| --- | --- |
| degree | count |
| 31 | 1 |
| 33 | 1 |
| 12 | 3 |
| 13 | 2 |
| 14 | 5 |
| 15 | 4 |
| 16 | 7 |
| 17 | 8 |
| 18 | 10 |
| 19 | 5 |
| 20 | 10 |
| 21 | 10 |
| 22 | 12 |
| 23 | 6 |
| 24 | 4 |
| 25 | 6 |
| 26 | 3 |
| 27 | 1 |
| 29 | 2 |

2. Gnm2:

|  |  |
| --- | --- |
| degree | count |
| 231 | 3 |
| 233 | 5 |
| 235 | 2 |
| 238 | 1 |
| 239 | 1 |
| 156 | 1 |
| 159 | 1 |
| 162 | 2 |
| 163 | 1 |
| 166 | 1 |
| 170 | 2 |
| 172 | 4 |
| 173 | 3 |
| 174 | 5 |
| 175 | 6 |
| 176 | 2 |
| 177 | 4 |
| 178 | 3 |
| 179 | 4 |
| 180 | 12 |
| 181 | 9 |
| 182 | 13 |
| 183 | 12 |
| 184 | 20 |
| 185 | 12 |
| 186 | 18 |
| 187 | 17 |
| 188 | 25 |
| 189 | 26 |
| 190 | 16 |
| 191 | 25 |
| 192 | 33 |
| 193 | 30 |
| 194 | 27 |
| 195 | 29 |
| 196 | 37 |
| 197 | 30 |
| 198 | 30 |
| 199 | 19 |
| 200 | 36 |
| 201 | 41 |
| 202 | 42 |
| 203 | 31 |
| 204 | 29 |
| 205 | 30 |
| 206 | 21 |
| 207 | 22 |
| 208 | 26 |
| 209 | 21 |
| 210 | 15 |
| 211 | 24 |
| 212 | 29 |
| 213 | 21 |
| 214 | 12 |
| 215 | 10 |
| 216 | 13 |
| 217 | 14 |
| 218 | 9 |
| 219 | 10 |
| 220 | 8 |
| 221 | 9 |
| 222 | 11 |
| 223 | 8 |
| 224 | 4 |
| 225 | 2 |
| 226 | 2 |
| 227 | 3 |
| 228 | 1 |
| 229 | 3 |
| 230 | 2 |

3. Gnp1:

|  |  |
| --- | --- |
| degree | count |
| 1 | 4 |
| 2 | 12 |
| 3 | 15 |
| 4 | 12 |
| 5 | 18 |
| 6 | 19 |
| 7 | 11 |
| 8 | 5 |
| 9 | 4 |

4. Gnp2:

|  |  |
| --- | --- |
| degree | count |
| 31 | 11 |
| 32 | 8 |
| 33 | 3 |
| 34 | 1 |
| 35 | 1 |
| 36 | 3 |
| 6 | 1 |
| 7 | 2 |
| 8 | 2 |
| 9 | 5 |
| 10 | 7 |
| 11 | 19 |
| 12 | 48 |
| 13 | 60 |
| 14 | 73 |
| 15 | 101 |
| 16 | 130 |
| 17 | 151 |
| 18 | 170 |
| 19 | 173 |
| 20 | 180 |
| 21 | 170 |
| 22 | 151 |
| 23 | 124 |
| 24 | 109 |
| 25 | 94 |
| 26 | 70 |
| 27 | 57 |
| 28 | 32 |
| 29 | 25 |
| 30 | 19 |

2. The results of the Stanford graphs are as follows:

|  |  |  |
| --- | --- | --- |
| **Graph** |  | **Scalefree** |
| amazon.graph.large | 1.325 | Yes |
| amazon.graph.small | 2.39 | Yes |
| dblp.graph.large | 1.31 | Yes |
| dblp.graph.small | 1.607 | Yes |
| youtube.graph.large | 1.560 | Yes |
| youtube.graph.small | 1.36 | Yes |

**Solution 2:**

* The rank of the nodes from highest to lowest closeness centrality is as follows:

F>C>H>D>B>E>A>G>I>J

* The Machine F and C will be the best Computers on the network to hold the data. These two nodes have the largest centrality closeness which means the shortest path between the other nodes to this node is the smallest. Therefore, if we choose these two nodes to hold the data, other machines will need fewest hops to access the data from these nodes compared to all others.

**Solution 3:**

The members that should be targeted to disrupt communication in the network are:

* Mohamed Atta
* Usman Bandukra
* Mamoun Darkazanli
* Essid Sami Ben Khemais
* Djamal Beghal
* Nawaf Alhazmi
* Raed Hijazi