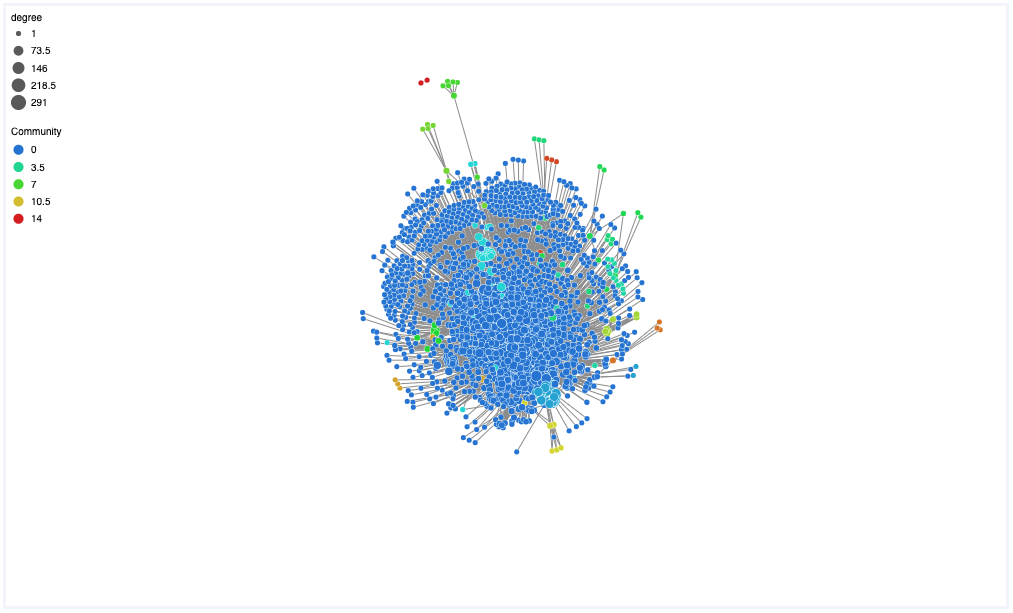
Greedy Community Detection Effects of Shifting Up HPO Tree

**Greedy Community Detection**: 14 Communities



Visualization using webweb of all 15 communities.

A close up of a map

Description automatically generated

Community 0 has 1,000+ nodes, that one is excluded here.

**Traversing up the HPO tree**

Mean Distance from HPO root: 6.55

A close up of a logo

Description automatically generated

Boxplot of HPO term distances from the root of the HPO tree (root being the node labeled ‘All’, HP:0000001). N=1485

A close up of a map

Description automatically generated

For community detection while changing shift distance the minimum distance threshold was held constant at 6 (the average distance from the root of the HPO tree). During community detection while changing minimum distance the shift distance was held constant at 3 (arbitrarily chosen).

Minimum distance is the minimum distance away from the root HPO term a node can be after shifting. In the event a node started at a distance closer to the root than minimum allowed, no shift occurred on that term and it’s original position was used. If shifting a node resulted in it being closer than the minimum distance the node was only shifted to the minimum distance. The intent behind this logic is to allow fringe nodes (HPO terms) to accumulate in more general and common nodes closer to the tree root, while preventing those already too generalized and close to the root from becoming overly generalized and loosing meaning.

**Adjusting the magnitude of the shift up the tree**

Holding the min distance at 6 and adjusting the shift variable has little to no effect on the number of the communities.

A close up of a map

Description automatically generated

Shift 1 Min 6, Excluding Largest

A close up of a map

Description automatically generated

Shift 2 Min 6, Excluding Largest

A close up of a map

Description automatically generated

Shift 3 Min 6, Excluding Largest

A close up of a map

Description automatically generated

Shift 4 Min 6, Excluding Largest

A screenshot of a computer

Description automatically generated

Shift 5 Min 6, Excluding Largest

A screenshot of a computer

Description automatically generated

Shift 6 Min 6, Excluding Largest

**Adjusting Min Distance Variable**

Adjusting the minimum distance allowed when shifting up the tree has a strong effect on the number of communities detected. The magnitude of the shift was held constant at 3. Moving from 4 to 5 min distance caused an explosion in the number of communities going from 5 to 10. At a minimum distance of 9 the number of communities leveled off at 14.

A screenshot of a cell phone

Description automatically generated

Shift 3 Min 1, Excluding Largest

A close up of a logo

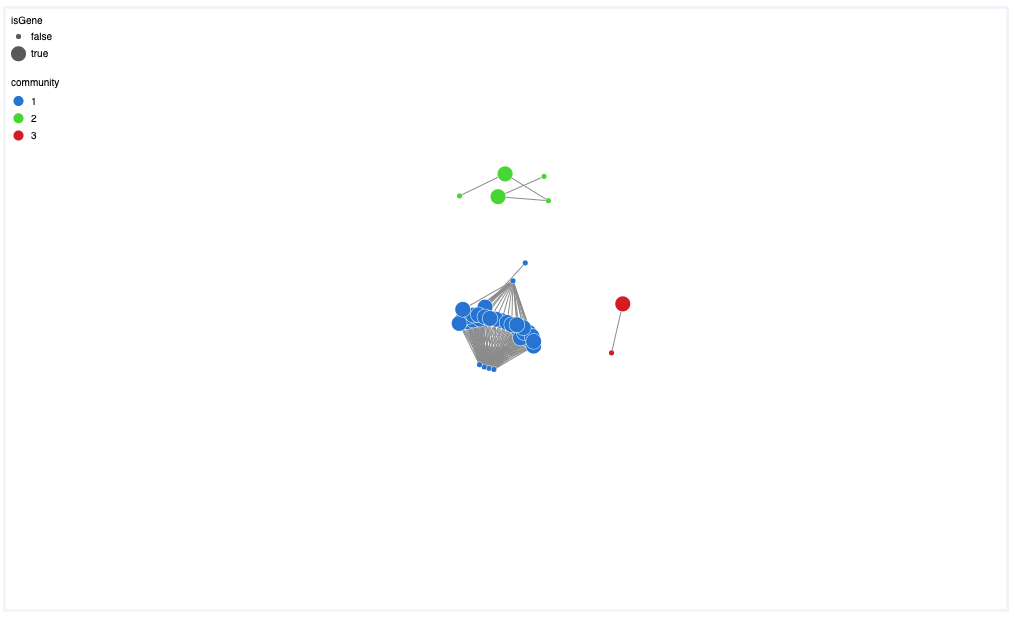
Description automatically generated

Shift 3 Min 2, Excluding Largest

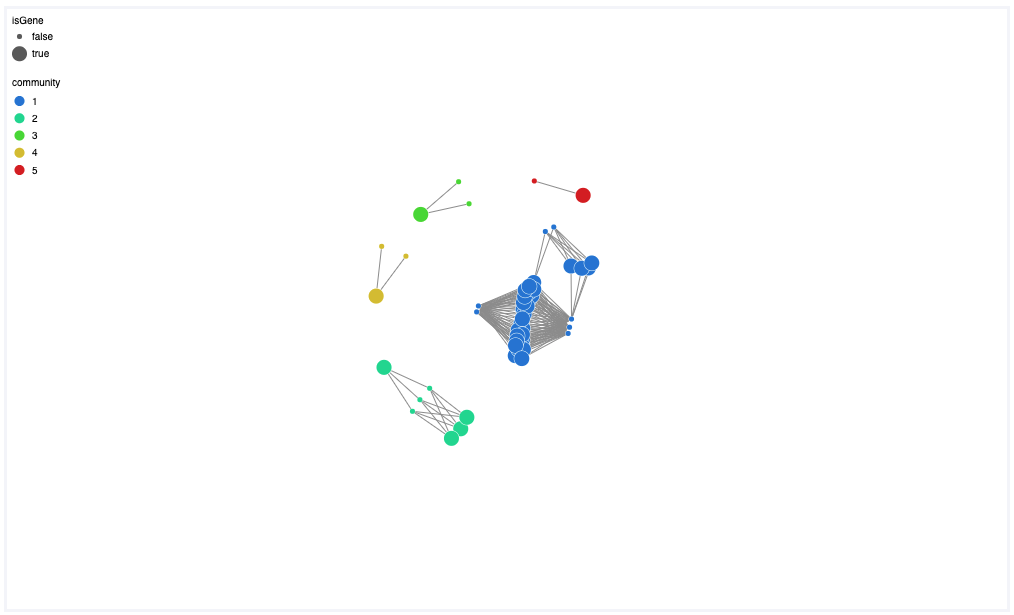
A picture containing screenshot

Description automatically generated

Shift 3 Min 3, Excluding Largest



Shift 3 Min 4, Excluding Largest



Shift 3 Min 5, Excluding Largest

A close up of a map

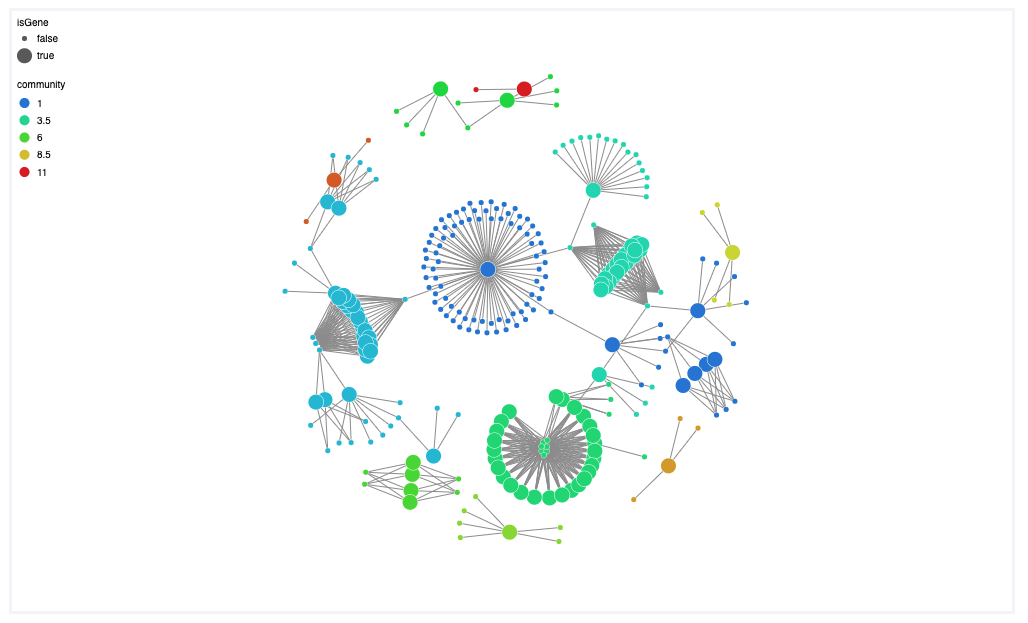
Description automatically generated

Shift 3 Min 6, Excluding Largest

A close up of a map

Description automatically generated

Shift 3 Min 7, Excluding Largest



Shift 3 Min 8, Excluding Largest

A close up of a map

Description automatically generated

Shift 3 Min 9, Excluding Largest

A close up of a map

Description automatically generated

Shift 3 Min 10, Excluding Largest

A close up of a map

Description automatically generated

Shift 3 Min 11, Excluding Largest