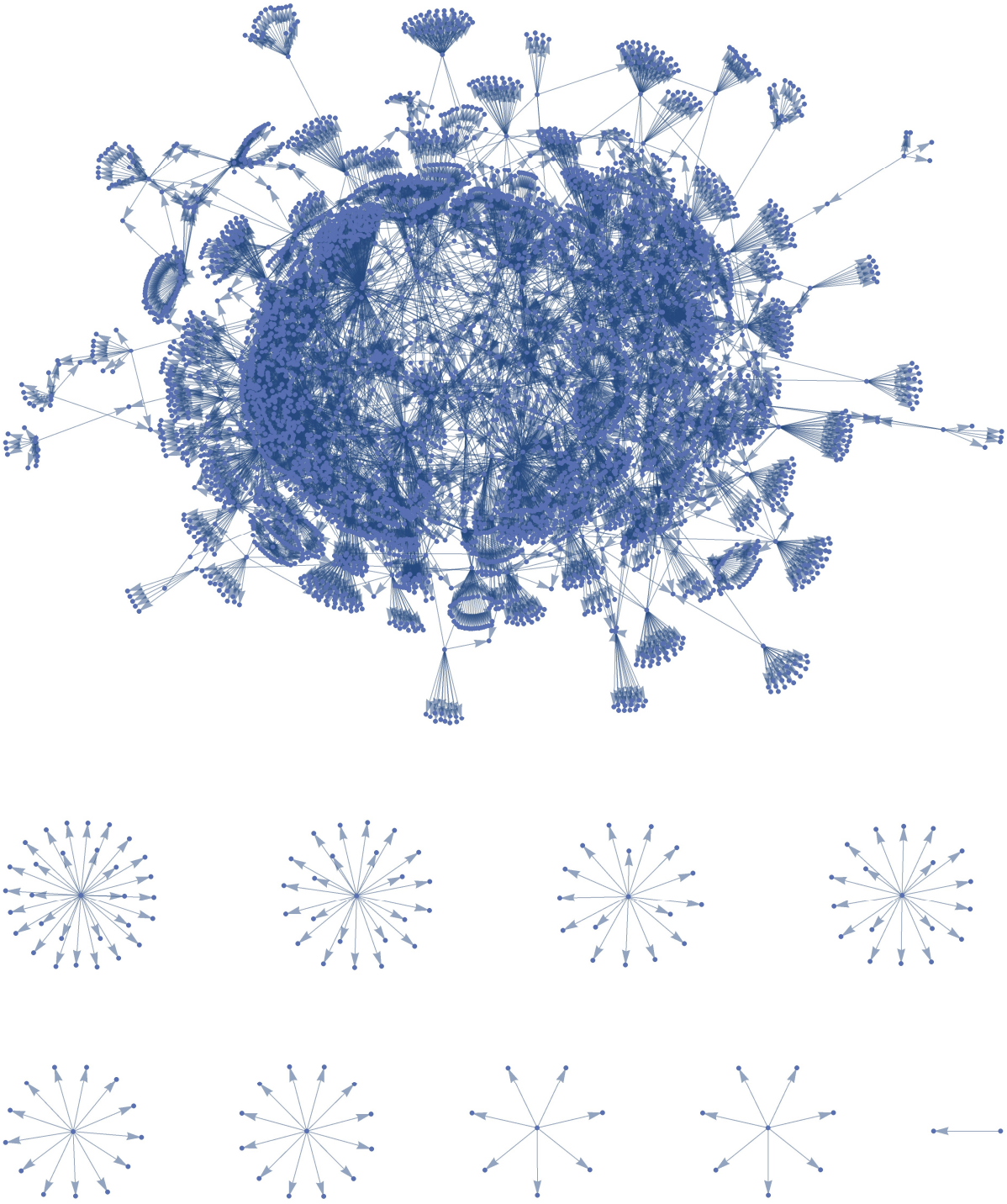
In[3]:=
```
AppendTo[$Path, FileNameJoin[{$HomeDirectory, "network-similarity"}]];
g = Import["test_int_id.gml", ImageSize → Full]
```

Out[4]=



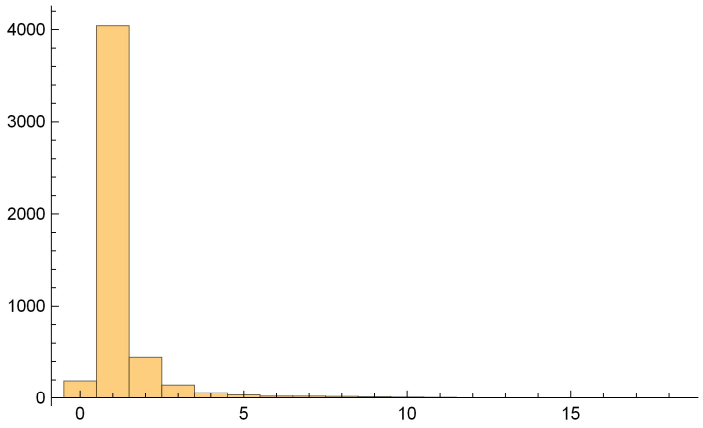
```
VertexCount[g]
EdgeCount[g]
```

```
4984
```

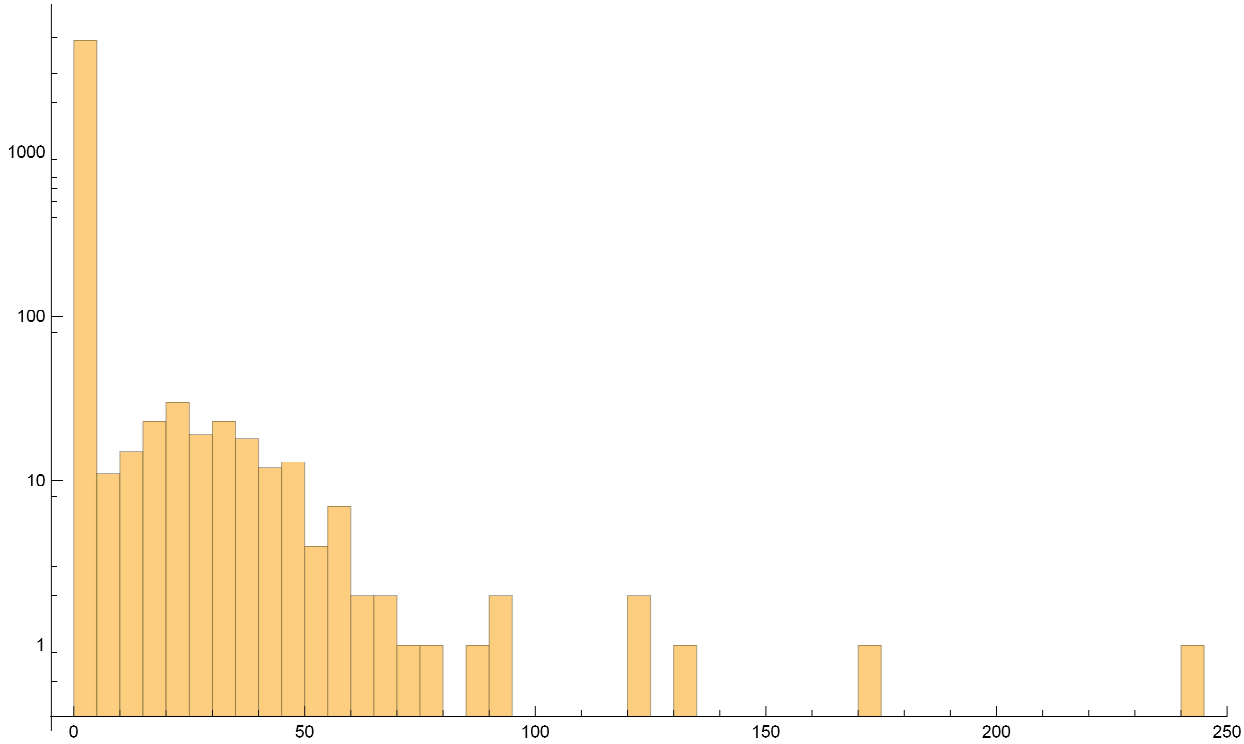
```
6403
```

```
HistogramList[VertexInDegree[g]][[2]]
Histogram[VertexInDegree[g], ImageSize → Medium]
HistogramList[VertexOutDegree[g], 50]
Histogram[VertexOutDegree[g], 50, ScalingFunctions → "Log", ImageSize → Full]
```

```
{187, 4043, 446, 143, 56, 33, 18, 16, 13, 10, 7, 3, 2, 1, 1, 1, 2, 1, 1}
```




```
{{0, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90,
  95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170,
  175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245},
 {4795, 11, 15, 23, 30, 19, 23, 18, 12, 13, 4, 7, 2, 2, 1, 1, 0, 1, 2, 0, 0, 0, 0,
  0, 2, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1}}
```



```
gUndirected = UndirectedGraph[g, Options[g]];
GraphAssortativity[gUndirected] // N
GraphAssortativity[gUndirected, "year"] // N
GraphAssortativity[gUndirected, "referenceCount"] // N
```

```
-0.36155
```

```
0.00134421
```

```
-0.0886864
```

```
comps = WeaklyConnectedComponents[g];
compsizes = {};
For[i = 1, i ≤ Length[comps], i++, AppendTo[compsizes, Length[comps[[i]]]]];
Print[Length[comps], " connected components with sizes ", compsizes]

outdegrees = VertexOutDegree[g];
For[i = 2, i ≤ Length[comps], i++,
 headLoc = Position[outdegrees[[comps[[i]]]], _?(# > 0 &)] // Flatten;
 head = comps[[i]][[headLoc]];
 Print[PropertyValue[{g, head[[1]]}, "title"], "\n\tOutdegree: ",
  outdegrees[[head[[1]]]], "\tYear: ", PropertyValue[{g, head[[1]]}, "year"]];
]
```

```
10 connected components with sizes {4848, 31, 24, 19, 17, 14, 13, 8, 8, 2}
A multi-seed dynamic local graph matching model for tracking
   of densely packed cells across unregistered microscopy image sequences
    Outdegree: 30    Year:
 2018
Robust Road Map Inference through Network Alignment of Trajectories
    Outdegree: 23    Year: 2018
Scheduling and Automatic Parallelization. Alain Darte, Yves Robert and
   Fr\ufffdd\ufffdric Vivien, Birkh\ufffduser, New York, ISBN 0-8176-4149-1
    Outdegree: 18    Year:
 2002
Answering Why-Not Questions on Structural Graph Clustering
    Outdegree: 16    Year: 2018
On the feasibility of precoding-based network alignment for three unicast sessions
    Outdegree: 13    Year: 2012
Question Answering Using Sentence Parsing and Semantic Network Matching
    Outdegree: 12    Year: 2005
Rahnuma: hypergraph-based tool for metabolic pathway prediction and network comparison
    Outdegree: 7    Year: 2009
Graph Matching Iris Image Blocks with Local Binary Pattern
    Outdegree: 7    Year: 2006
Hadamard equivalence via graph isomorphism
    Outdegree: 1    Year: 1979
```

What percentage is in the giant component? What is the diameter?

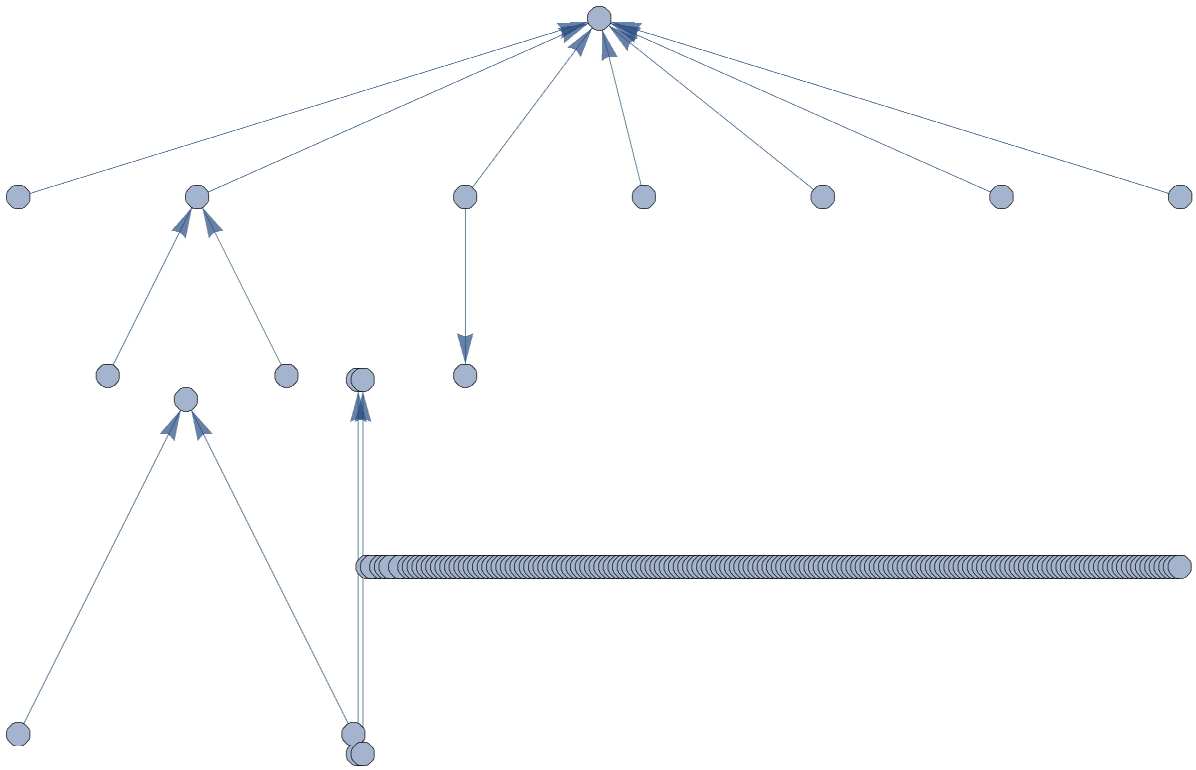
```
gCC = Subgraph[g, comps[[1]]];
Length[VertexList[gCC]] / Length[VertexList[g]] // N
GraphDiameter[UndirectedGraph[gCC]]
```

```
0.972713
```

```
12
```

Also investigate the subgraph of just the parents.

```
outdegrees = VertexOutDegree[g];
indegrees = VertexInDegree[g];
parents = Position[outdegrees, _?(# > 0 &)] // Flatten;
Subgraph[g, parents, ImageSize → Full, GraphLayout → "LayeredDrawing"]
For[i = 1, i ≤ Length[parents], i++,
 If[indegrees[[parents[[i]]]] > 0,
  Print[PropertyValue[{g, parents[[i]]}, "title"]];
  Print["\tIndegree ", indegrees[[parents[[i]]]],
   ", citation count ", PropertyValue[{g, parents[[i]]}, "citationCount"]]
 ]
]
```



```
Fifty years of graph matching, network alignment and network comparison

    Indegree 2, citation count 28

A linear programming approach for the weighted graph matching problem

    Indegree 7, citation count 113

Conceptual graph matching: a flexible algorithm and experiments

    Indegree 1, citation count 9

Detecting Self-mutating Malware Using Control-Flow Graph Matching

    Indegree 1, citation count 40

A new algorithm for inexact graph matching

    Indegree 1, citation count 12

BIG-ALIGN: Fast Bipartite Graph Alignment

    Indegree 2, citation count 16
```

Let's look at the titles for our most commonly cited (within the network) and highest outdegree nodes.

```
mostCited = Part[VertexList[g], Ordering[DegreeCentrality[g, "In"], All, Greater]];
For[i = 1, i ≤ 10, i++,
 Print[PropertyValue[{g, mostCited[[i]]}, "title"]];
 Print["\tIndegree ", indegrees[[mostCited[[i]]]],
  ", citation count ", PropertyValue[{g, mostCited[[i]]}, "citationCount"]]
]
```

```
An Algorithm for Subgraph Isomorphism

    Indegree 18, citation count 760

THIRTY YEARS OF GRAPH MATCHING IN PATTERN RECOGNITION

    Indegree 17, citation count 468

The Hungarian method for the assignment problem

    Indegree 16, citation count 2843

A graduated assignment algorithm for graph matching

    Indegree 16, citation count 504

An eigendecomposition approach to weighted graph matching problems

    Indegree 15, citation count 321

A distance measure between attributed relational graphs for pattern recognition

    Indegree 14, citation count 300

Pairwise Global Alignment of Protein Interaction Networks by Matching Neighborhood Topology

    Indegree 13, citation count 84

Computers and Intractability: A Guide to the
  Theory of NP-Completeness (Michael R. Garey and David S. Johnson)

    Indegree 12, citation count 6

Topological network alignment uncovers biological function and phylogeny

    Indegree 12, citation count 120

Network Motifs: Simple Building Blocks of Complex Networks

    Indegree 11, citation count 2730
```

```
mostCiting = Part[VertexList[g], Ordering[DegreeCentrality[g, "Out"], All, Greater]];
For[i = 1, i ≤ 10, i++,
 Print["Outdegree ", outdegrees[[mostCiting[[i]]]],
  ": ", PropertyValue[{g, mostCiting[[i]]}, "title"]]
]
```

```
Outdegree 241: Networks for systems biology: conceptual connection of data and function

Outdegree 170: THIRTY YEARS OF GRAPH MATCHING IN PATTERN RECOGNITION

Outdegree 134: The graph matching problem

Outdegree 122: Fifty years of graph matching, network alignment and network comparison

Outdegree 120: Methods and Implementations of Road-Network Matching

Outdegree 93: Graph-based methods for analysing networks in cell biology

Outdegree 90: Efficient Graph Matching Algorithms

Outdegree 85: Software verification and graph
   similarity for automated evaluation of students\u2019 assignments

Outdegree 76: Multiple Alignment, Communication Cost, and Graph Matching

Outdegree 73: Modeling cellular machinery through biological network comparison
```