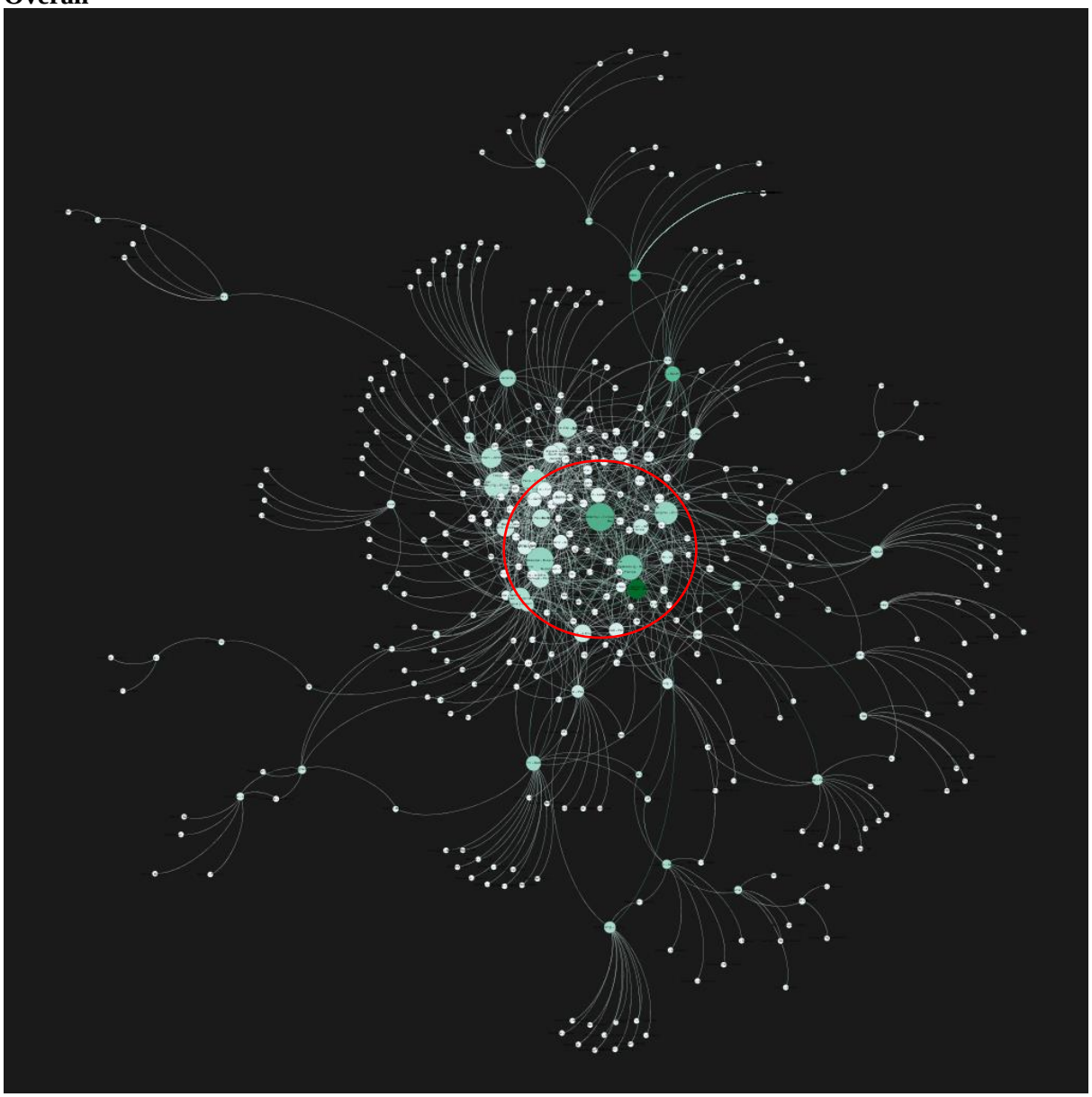
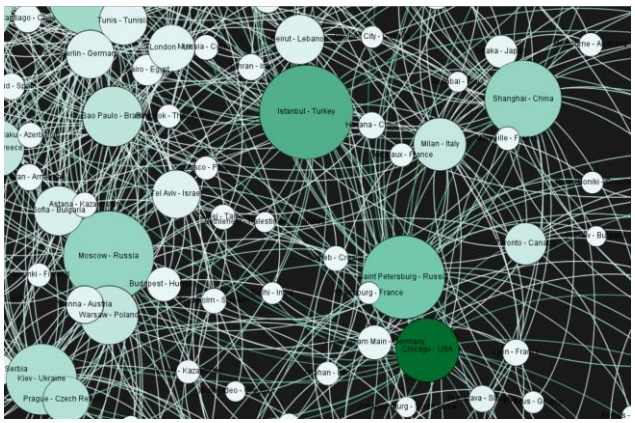
Picture of Graph 1
Overall



Locally zoomed in



Picture of Graph 1
Graph Drawing Contest 2009: See http://www.graphdrawing.de/contest2009/gdcontest2009.html
Background: City Mystery Graph
This is a graph with 391 nodes and 911 undirected edges. The nodes represent cities. The 17 cities of the past Graph Drawing Conferences are fixed, while the remaining cities are free to be moved. There is no requirement concerning edge shapes or layout style.

Description:
A: how to create the pictures. The raw file is in text format containing the nodes(city) and edges (city to city connections). Data cleaned (remove the not required data) and pre-processed into csv format. The picture is created using Gephi (*Gephi* 0.9.1). The algorithm used is Gephi's built in Force Atlas. The Force Atlas layout algorithm comes under a category of algorithms called force-directed algorithms. There's usually a trade-off between quality and speed when it comes to graph layout algorithms. Force Atlas emphasizes the former over the latter; that is, the Force Atlas layout algorithm gives more weight to the quality of the layout than the speed with which it has been computed. (*Force atlas* 2017)

During the creating of the graph combination of automatic layouts and manual adjustment are used, the automatic layouts are for initial graph generating, while manual adjustments are for those nodes being repulsed too far away.

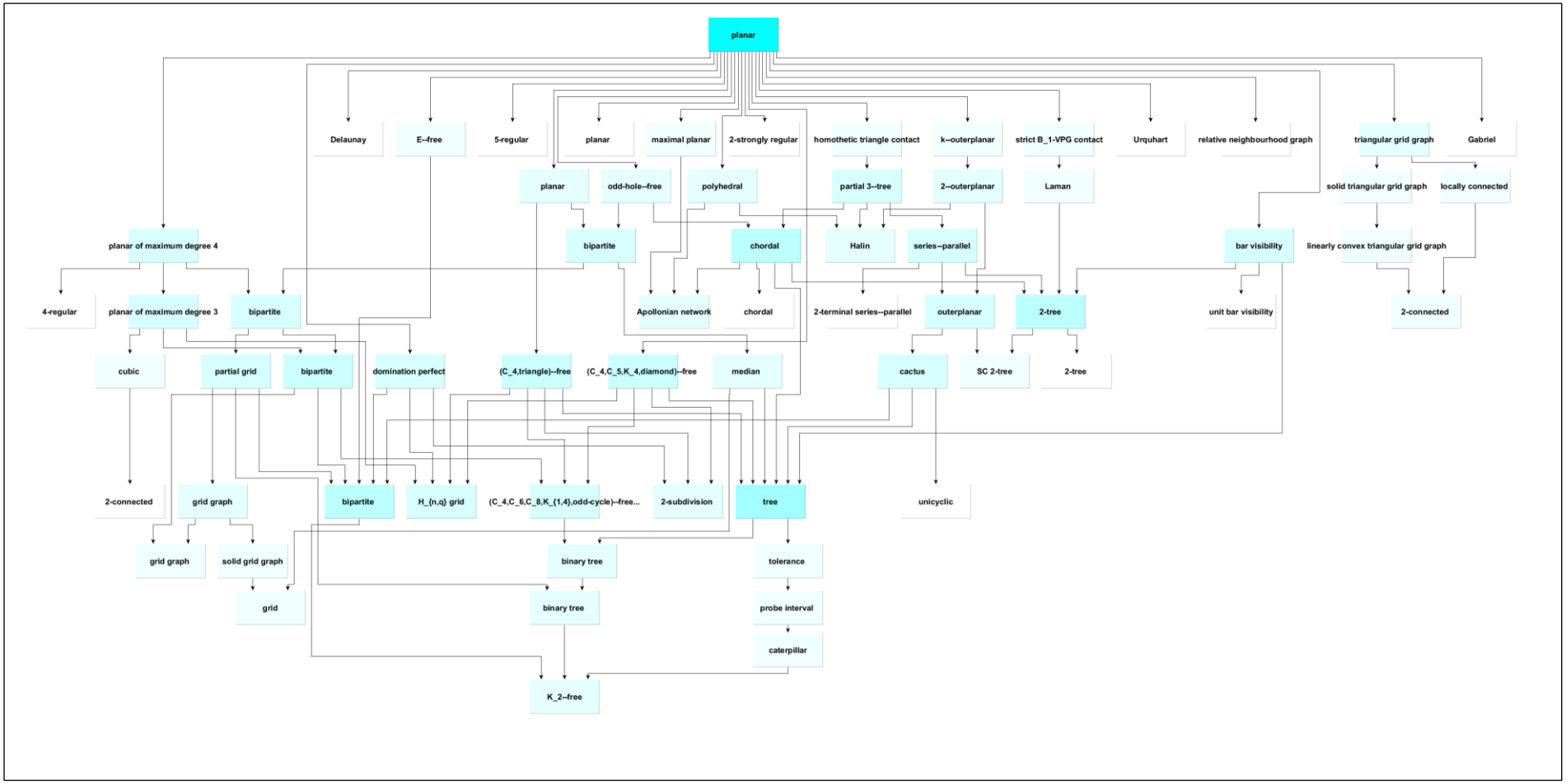
After the analysis Istanbul is having highest degree (Number of nodes it is connected to), and Chicago is having highest betweenness centrality (number of nodes it could possibly be connected with)

id	label	degree
46	Istanbul - Turkey	50

B. self-evaluation: strength and weakness
Strength: The most important and critical information being properly visualised. The nodes, with high degree (Degree 2017), are showing large proportion of size visually appealing. *Between Centrality (Between Centrality Wikipedia* 2017) is being colour coded showing darkest colour with deep green, the least central ones with light colour. The background has been set as black, by doing this the least connected node (light colour) can also be shown properly on the graph.
The edges being coloured followed the between centrality colour setting, by doing this will improve the visualisation for between centrality.

Weakness:
There are still edges crossing between the Nodes, Edge bunding (Edge Bundling, Holten 2006) could be introduced to reduce Visual Complexity. The advantage of setting black as background colour, at the meantime, introducing some issue for the labelling, because the labelling is set as black.

Picture of Graph 2



Picture of Graph 2
Graph Drawing Contest 2015: See http://www.csun.edu/gd2015/topics.htm
Background: Graph Classes
This graph models a set of well-known subclasses of planar graphs and their relations. The graph has 65 vertices and 101 edges. The resulting layout of the graph should contain the label of the vertices, provided as the description of the nodes, and should give a good overview on the hierarchy of the graph classes.

Description:
A: how to create the pictures. The tool used for creating the graph is yeD (yeD 3.17). The raw data file is presented in the GraphML File Format. In order to process in yeD, data pre-processing is done, the file being transformed to excel format with Node and Edge properties.

After analysis of the graph. the relationship between the nodes are hierarchy, the planar is on the top, K_2-free being the bottom.

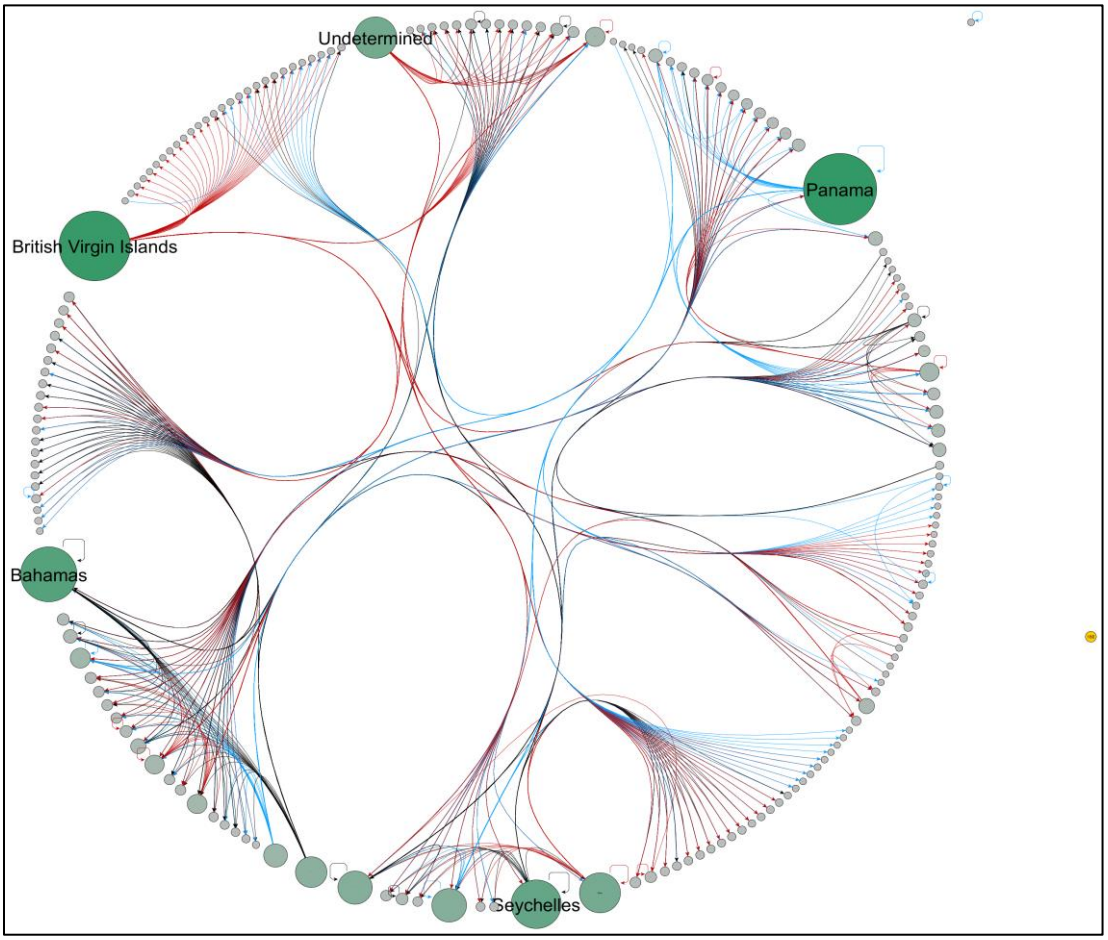
The picture layout is using yeD's hierarchical layout due to nature of the relationship. Property Mapper being used to map the Node's label, so each node could have the name shown on the graph. Customised Palette created to present the graph as the intended way. Colour coding also used by using the Centrality Measures function in yeD. The colour is presented with most number of connected node showing dark colour and the least connected node showing light colour. This turns out the Planar is having highest number of edges and tree ranked second in terms of number of edges

B. self-evaluation: strength and weakness
Strength: During the implementation, the edges pointing upward have been successfully avoided, the nodes plotted evenly, Edges have been kept straight and vertical where it is possible. In terms of space utilisation, I also managed to maximise the available space.
One extra highlight is using colour coding to show the number of edges for each node given extra information when visualising.

Weakness: A few long edges plotted, for examples the Planar to its children. This could be improved by increase the Planar node size. Edges crossing are still happening can further reduced with more optimisation

Possible better way: In this exercise tree layered drawing is used, alternatively radial drawing could be used, using planar as the centre, this will help to reduce the edges crossing. Node size can be encoded to distance of the parent (Planar) to give extra information of the graph

Picture of Graph 3



Picture of Graph 3

Graph Drawing Contest 2016: See http://www.graphdrawing.de/contest2016/topics.html

Background: Panama Papers

This graph models the relations between countries through off-shore companies as revealed by the so-called "Panama papers". A database to create a weighted directed graph that shows the relationships between countries. A directed edge from country A to country B with weight w means that there are w Offshore Entities in country B that are linked to a company in country A.

Description:

A: how to create the pictures. The tool used for creating the graph is yeD (yeD 3.17). The raw data file is presented in the XML format, with each country with offshore entities, an adjacency of its incoming edges with their weight have been created. In order to process in yeD, data pre-processing is done, the file being transformed to excel format with Node and Edge properties, weighted properties also included in the edge properties

The picture layout is using yeD's circular layout, this circular layout giving good symmetry of the graph, each country being evenly distributed around the circle. Weighted connected edges considered for both incoming and outgoing edges. This is to identity the highest weighted edges. It turns out British Virgin Islands, Panama and Bahamas are the top 3 countries. Property Mapper being used to mapping each country's name to the node's label, so each node could have the name shown on the graph. The colour coding applied for each node, with darkest colour showing high weighted nodes and lightest colour showing least weighted nodes

B. self-evaluation: strength and weakness

Strength: Visualization complexity reduced by successfully implemented edge bunding (*Edge Bundling*, Holten 2006). Edges arrows indicated the incoming and out coming edges

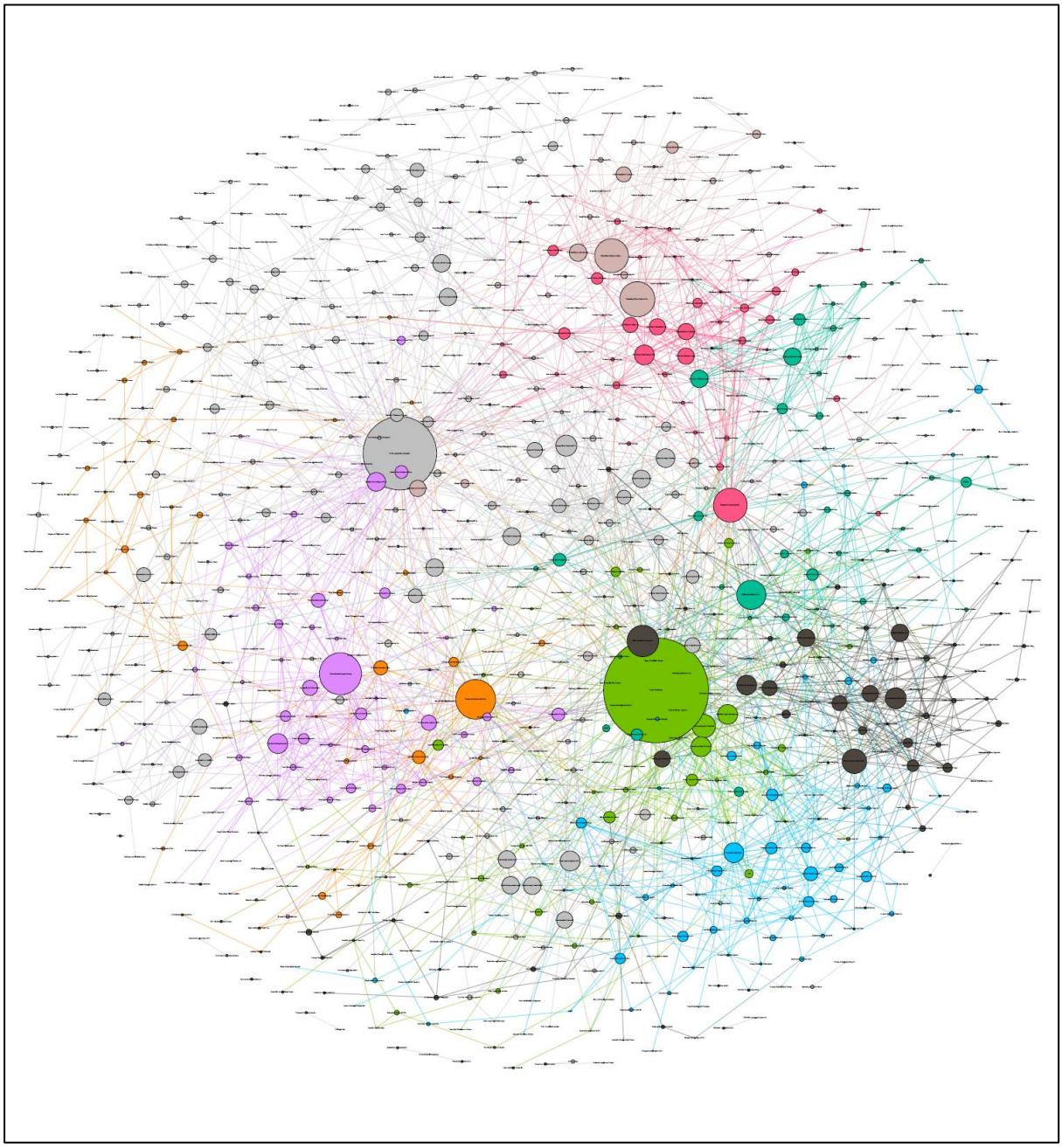
The circle structure has the advantage maximise the space usage.

One more highlight is using colour coding to show the number of edges for each node given extra information when visualising.

Weakness: Currently the node size and node colour showing the same information, weighted connected edges. This can be improved by coding of them to see only incoming or outgoing edges. In order to have the important node showing obviously, the less important node's font size being sacrificed

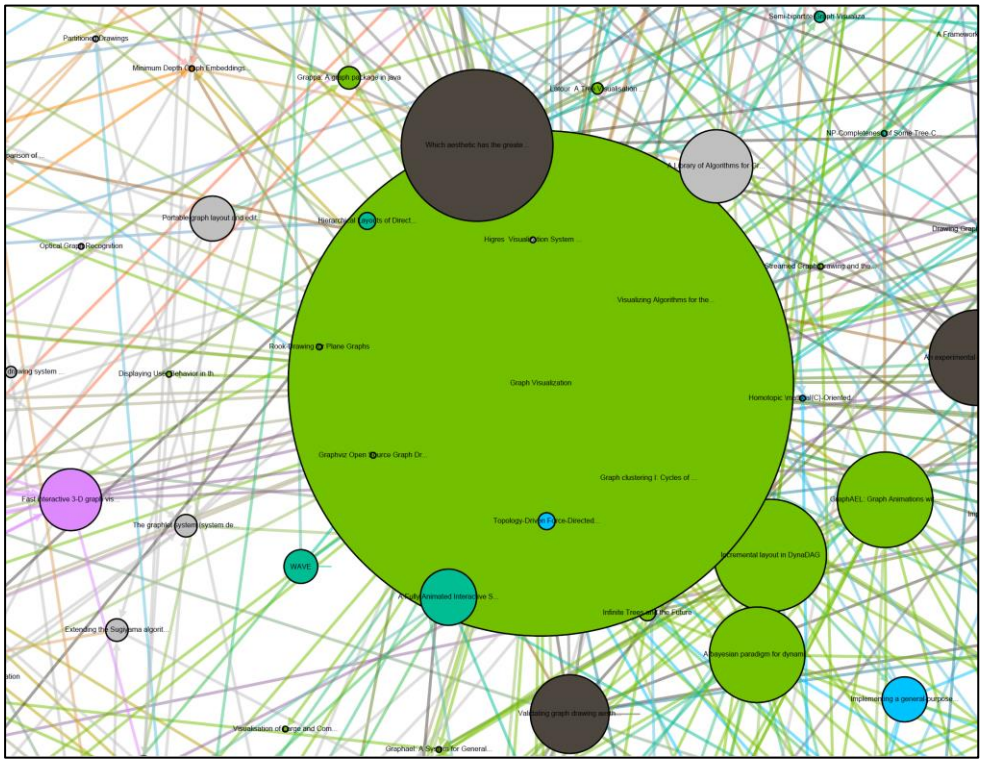
Possible better way: The weighted edges can be coded with size of edges to give better visualisation of the weight information.

Picture of Graph 4
Overall

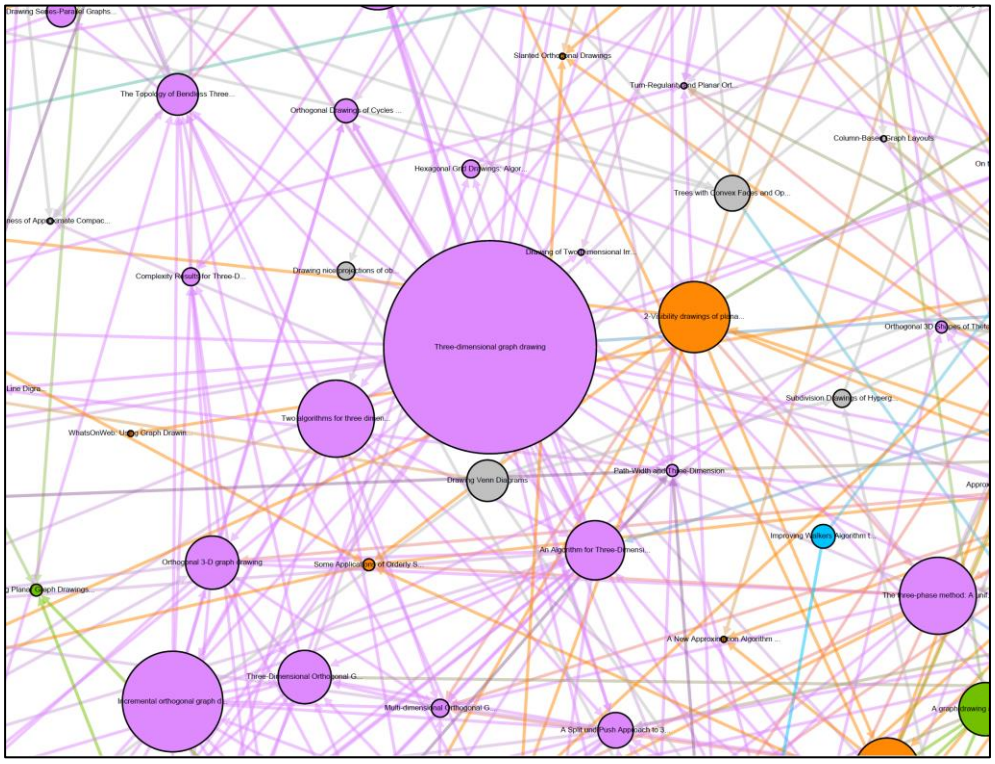


Picture 4

Locally zoomed in



Locally zoomed in



Picture of Graph 4
Graph Drawing Contest 2017: See: http://www.graphdrawing.de/contest2017/topics.html
Background: Graph Drawing Citations
We extracted data of all publications in the Proceedings of Graph Drawing between 1994 and 2015. For each paper, we provide the title, the authors, the institutions of the authors, the year, and the citations. Your task is to create a nice visualization of the given citation network. It is desirable but not compulsory to make use of the extra data (authors, institutions, years) or a subset of it to explore interesting structures in the network.

Description:
A: how to create the pictures. The tool used for creating the graph is Gephi (*Gephi* 0.9.1). The raw data file is presented in the XML format (with 15632 rows of data). Data pre-processed being done, there are 151 nodes do not have any cites or cited work, they are isolated nodes. These nodes are removed to improve the visualisation. After removing these isolated nodes there are 1053 valid nodes available. Two csv files (Node and Edge list) created for Gephi to import.

Number of out-degree means the paper being cited, and number of in-degree means the paper cites.

After analysing the data, paper *Graph Visualization*, published in 2011, being cited most for all 1053 papers, total 89 times being cited, interestingly, this paper itself did not cite any other paper. Paper *Selected Open Problems in Graph Drawing*, published in 2003, cites 24 papers and being cited 10 times.

Modularity class (*Modularity class* 2017) used to colour coding the group, further filtering has been applied on giant component, the number of communities reduced from 34 to 14, this will help to reduce the ambiguousness for modularity class visualization

The layout algorithm used is Gephi's built in layout Fruchterman & Reingold (*Force-directed graph drawing*, 2017) combination with manual adjustment.

B. self-evaluation: strength and weakness
Strength: There are a few highlights about the graph thanks to the force-directed-graph, the overall nodes are distributed evenly with similar edge length. The graph is in great symmetry format. Each edge being drawn with arrow, showing the direction of the connection. Edge being coloured and Modularity both being coloured based on modularity class this give the clear picture of structure of the network. Last but not the least, out-degree number are encoded with node size given the highlight of the most cited paper

Weakness:
The number of edge crossing can be reduced. With help with edge bunding improved visualization is expected to achieve. Because Node size is used to show the number of paper being citied, some nodes being masked by large node.
Further modularity analysis can be done to improve the visualisation showing the inter connection between the papers

Possible better way: In this exercise 2D layered drawing is used, alternatively 2.5 layer can be used, this will help to reduce the edges crossing. Edge bunding method can be used to reduce the visual complexity

References

1. *Force atlas 2017*
   Using the Force Atlas layout algorithm viewed on 8/9/2017
   https://www.packtpub.com/mapt/book/big_data_and_business_intelligence/9781783987405/3/ch03lvl1sec36/using-the-force-atlas-layout-algorithm
2. *Degree* 2017
   Degree (graph theory) Wiki
   https://en.wikipedia.org/wiki/Degree_(graph_theory)
3. *Modularity class* 2017
   Modularity_(networks) Wiki
   https://en.wikipedia.org/wiki/Modularity_(networks)
4. *yeD* 3.17
   http://www.yWorks.com.
5. *Gephi* 0.9.1
   https://gephi.org/users/download/
6. *Force-directed graph drawing*, 2017
   https://en.wikipedia.org/wiki/Force-directed_graph_drawing
7. *Edge Bundling*, Holten 2006
   http://www.aviz.fr/wiki/uploads/Teaching2014/bundles_infovis.pdf
8. *Between Centrality Wikipedia* 2017
   Between Centrality Wikipedia page viewed on 8/9/2017)
   https://en.wikipedia.org/wiki/Betweenness_centrality