A Novel Centrality Based Method for Visual Analytics of Small-World Networks

Chun-Cheng Lin · Weidong Huang · Wan-Yu Liu*· Sheng-Feng Wu

Abstract Nowadays the network data that we need to deal with and make sense of is becoming increasingly large and complex. Small-world networks are a type of complex networks whose underling graphs has small diameter, shorter average path length between nodes, and a high degree of clustering structures; and can be found in a wide range of scientific fields, including social networks, sociology, computer science, business intelligence and biology. However, conventional visualization algorithms for small-work networks lead to a uniform clump of nodes or are restricted to a tree structure, making the network structure difficult to identify and analyze. This work provides a new visual analytical method to improve the situation. Different from previous methods based on spanning trees, this method first generates a weighted planar sub-network based on the measurement of network centrality metrics. A force-directed algorithm based on node-edge repulsion is then applied to visualize this sub-network into a proper layout for better understanding of the data. Finally, the remaining links are placed back to maintain the original network's integrity. The experimental results show that compared to previous methods, the proposed method can be more effective in differentiating clusters and revealing relationship patterns among individual nodes and clusters in the network. Furthermore, the proposed method is applied to a data of the semiconductor wafer manufacturing industry as a case study. The work shows that this new approach allows users to gain useful insights into the data.

Keywords Small-world network · network analysis · visual analysis · information visualization · force-directed method

1. Introduction

A network (or graph) consists of nodes and links (or edges). Each node in a network represents an entity such as a person, an organization, or a country. A link represents a relationship between the entities that could be communication, collaboration, friendship, trade, or any other types of relationships. A small-world network is a special type of networks in which most nodes can reach the other nodes through a small number of hops. Hence, a small-world network is generally a graph with small diameter, shorter average path length between nodes, and a high degree of clustering structures. Compared to random networks, the main feature of small-world networks is that the majority of nodes do not mutually connect. However, a small number of links can connect the entire network and provide higher cluster characteristics (Watts and Strogatz 1998; Yang and Chen 2018). In addition to social networks, this characteristic graph also appears in sociology, computer science, and biology (Bassett and Bullmore 2017; Das et al. 2017).

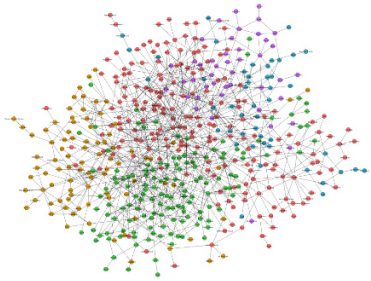
Nowadays the network datasets that we need to deal with are becoming increasingly large and complex. Hence, how to visually analyze and make sense of them have been one of important research topics. When it comes to network visualization, a key task is to ensure that the nodes and links have an appropriate geometric layout so that the visualization looks visually pleasing and is easy to understand (Huang et al. 2013; Huang et al. 2016). The following are some of the aesthetic criteria that have been widely used to judge the quality of network visualizations, and it is generally accepted that a network visualization will be effective if it conforms to these criteria (Bender-deMoll and McFarland 2006; Bertin 1983; Bi et al. 2019; Brandes et al. 2012; Lin et al. 2018): same link length, uniform node distribution, minimization of link crossings, and maximization of symmetries.

C.-C. Lin · S.-F. Wu
Department of Industrial Engineering and Management, National Chiao Tung University, Hsinchu, Taiwan
E-mails: cclin321@nctu.edu.tw, a860318@gmail.com

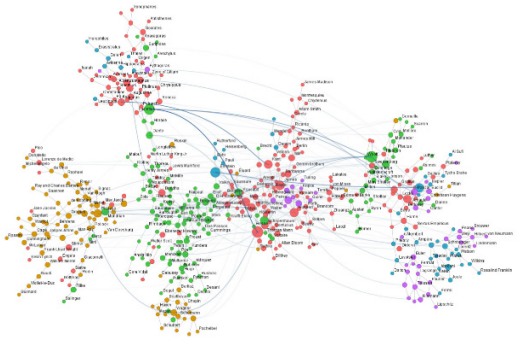
W. Huang
Faculty of Transdisciplinary Innovation, University of Technology Sydney, Ultimo, NSW 2007, Australia
E-mail: weidong.huang@uts.edu.au

* W.-Y. Liu (Corresponding author)
Department of Forestry, National Chung Hsing University, Taichung, Taiwan;
Innovation and Development Center of Sustainable Agriculture, National Chung Hsing University, Taichung, Taiwan
E-mail: wyliu@nchu.edu.tw

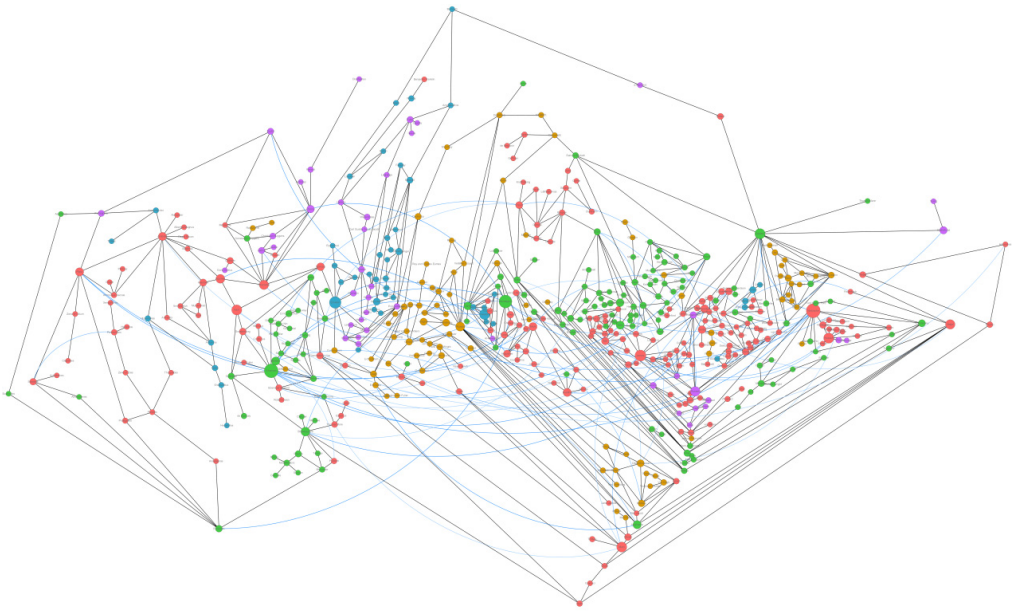
However, because small-world networks possess a special structure, a lot of conventional visualization methods that aim to meet the above aesthetic criteria often cause nodes and links to mutually overlap, making the identification of the network structure difficult. Force-directed graph drawing methods (Eades 1984) are commonly used to visualize networks, due to the fact that this kind of methods are easy to program in code and that the resulting visualizations are generally suitable for random and symmetric networks. However, the use of these methods on small-world networks can cause excessively even distribution of nodes and an excessive number of link overlaps, making it difficult to find paths between nodes or identify relationship patterns in the network (Gibson and Vickers 2016; van Ham and Wattenberg 2008). Fig. 1(a) shows an example visualization of a conventional force-directed algorithm, in which some simple clustering information can be observed, but almost all the link information is hidden.



(a) Conventional force-directed graph drawing method



(b) The previous method based on a minimum spanning tree (van Ham and Wattenberg 2008)



(c) The proposed method based on a planar subgraph, in which the links between clusters and individual nodes can be easily identified

Fig. 1. Comparison of the drawings of a small-world network of well-known individuals in history (Love 2007) with 515 nodes and 1070 links using different methods, in which each node represents an individual, while color indicates profession, including artists (beige), philosophers (red), writers (green), scientists (blue), and mathematicians (purple).

This work aims to propose and evaluate a novel visual analytic approach specifically designed for small-world networks for sense making and decision making purposes. The proposed method first assigns each link of the network graph a different weight according to its importance, and then finds a maximum-weight planar subgraph behind the graph. Note that the subgraph has the same number of nodes with the original graph, but has a smaller number of links. Then, while the graph planarity, the remaining links are added back as curves with a different color in the drawing so that unnecessary link crossings are avoided. By doing so, the backbone of the graph with more importance can be visualized. To ensure that the backbone will lead to meaningful visualization results, whenever possible, setting weights should follow the following constraints (van Ham and Wattenberg 2008): if the original network is connected, the resulting drawing should be connected as well; while some smaller-weight links are eliminated for the purpose of better visualization, the number of links that are to be deleted should be kept to a minimum; the resulting weighted graph drawing should maintain the clustering structure of the original network.

To meet the aforementioned constraints in setting weights, graph *centrality* metrics have been used in prior research. *Centrality* was first proposed for social network analysis to reflect the sociology origin of nodes, but it has been widely used to measure the importance of nodes and links in various fields (Bi et al. 2018; Freeman 1979; Opsahl et al. 2010), including analysis of small-world networks (Abbasi et al. 2012; Gómez et al. 2013; Sohn and Kim 2010; van Ham and Wattenberg 2008; Zuo et al. 2012). Among others, the method proposed in (van Ham and Wattenberg 2008) works as follows: first, a minimum-cost (i.e., maximum-weight) spanning tree is produced based on betweenness centrality values of edges. Next, links are placed back to the tree as long as they do not break the planarity of the updated subgraph. A conventional force-directed method is then applied to the final subgraph to generate a planar drawing. Finally, the rest of the links are added back to the drawing to produce the final visualization (Fig. 1(b)).

The method in (van Ham and Wattenberg 2008) is based on a minimum-cost spanning tree, and hence the resulting drawing is restricted to this tree structure (see Fig. 1(b)). Instead of a tree, the proposed method is based on a maximum-weight planar subgraph so that the drawing will not be limited by the tree structure, be better conform to the original graph structure, and conform to the aforementioned conditions. Note that the maximum-weight planar subgraph concerned in this work has the same number of vertices with the minimum-cost spanning tree concerned in van (Ham and Wattenberg 2008), but has more node-node-connection information, which can provide more insights into the concerned network. Once the planar graph is expended with the maximum possible links, an improved force-directed method based on node-edge repulsion (Bertault 2000) is used to lay out the subgraph, in which the rest of links are then added back and drawn as curves to produce the final drawing. See a comparison of drawings produced by the method in (van Ham and Wattenberg 2008) (Fig. 1(b)) and the method proposed in this work (Fig. 1(c)).

Note that each of Figs. 1(a), 1(b), and 1(c) includes node label text. The size of node label text in each figure differs because Figs. 1(a) and 1(c) are generated by our implementation, but Fig. 1(b) is form the previous work (van Ham and Wattenberg 2008), as we did not implement their method. In fact, this work focuses on drawing the graph structure, not on the vertex labeling. Vertex labeling is provided only for readers to understand the connection information.

In addition to the improved visualization method for small-world networks, this work also demonstrates the practical value of graph centrality measurements in visualization of small-world networks. The experiment results show that the drawings produced by the proposed method can show the relationships between nodes clearly and at the same time maintain the structure of the original network. Furthermore, the use of an improved force-directed method makes the planar drawing of the subgraph more visually pleasing, which is because the improved force-directed method was designed to avoid edge overlapping as much as possible. This in turn significantly improves the defect of node and link overlapping which is often found in conventional force-directed methods. In addition, for analysis of the manufacturing data of semiconductor wafers, statistical methods are often used, which often results in complex data information being presented in separate forms (Lee and Kim 2014). Therefore, this work conducts a case study by applying the proposed method in semiconductor wafer data to demonstrate its benefits.

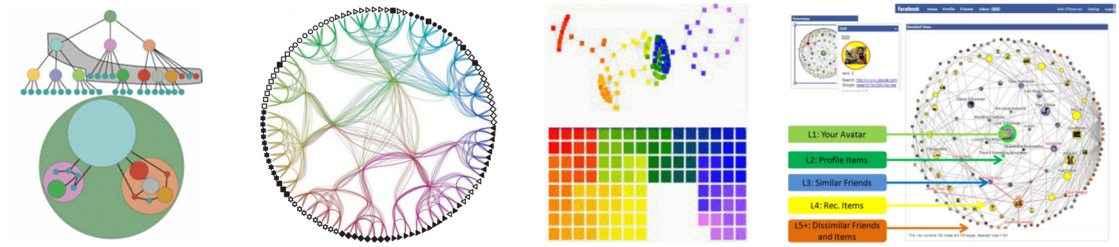
2. Literature Review

2.1 Network analysis

The use of social media has been popular among users. Social networks formed by those users contain rich information, and it is important to analyze them to retrieve those information for various purposes. For example, relationships between individuals in a network is important as it can help analysts understand which individuals within a network have influence and how individuals behave to form relationship to exert

influence (Borgatti et al. 2009). Furthermore, previous studies have also conducted cluster analysis on the graph structure of a social network (Burger et al. 2013; Guan and Yuen 2013). A network can contain a lot of clusters. Because similar background environment or characteristics can increase interaction between nodes and concentrate the nodes, clusters usually have the following characteristic: links in the same cluster are tighter (stronger interaction between nodes within the cluster) and links between different clusters are sparser (less interaction between two nodes respectively from different clusters).

Visualization has been used to help with network analysis. A lot of works have also proposed different visualization methods. Archambault *et al.* (Archambault et al. 2011) proposed the TugGraph (Fig. 2(a)) to effectively preserve the hierarchy structure when browsing network paths. Jia *et al*. (Jia et al. 2011) proposed a hierarchical edge bundling method (HEB) (Fig. 2(b)) to allow graphs to be visualized with better modularization characteristics while maintaining the overall balance. Wong *et al.* (Wong et al. 2012) proposed GreeCurve (Fig. 2(c)) for graphs with small-world characteristics. This method mainly uses a fractal-based design and spatial clues to group similar nodes together, eliminating physical links between nodes in visualization. Other than producing visually compelling and perceptionally effective visualizations, interaction is also another aspect of research for big network data. For example, Gretarsson *et al.* (Gretarsson et al. 2010) developed a set of visualization tool called SmallWorlds (Fig. 2(d)), which provides good recommendation mechanisms based on Facebook's user data and interactions between users. This tool also uses an outstanding visualization interface to display recommendation results.



(a) (b) (c) (d)

Fig. 2. Some notable network analysis tools: (a) TugGraph (Archambault et al. 2011) (b) HEB (Jia et al. 2011) (c) GreenCurve(Wong et al. 2012) (d) SmallWorlds (Gretarsson et al. 2010).

2.2 Network centrality

In graph theory and network analysis, centrality is an indicator for finding important nodes and links (Bhandari et al. 2017; Bonchi et al. 2016; Freeman 1979; Opsahl et al. 2010). Centrality-related studies have been conducted in a wide range of scientific fields, such as sociology and biology (Abbasi et al. 2012; Zuo et al. 2012). What kind of influence a node has on a network and what central position it occupies are among the earliest concepts that were explored in network analysis. Centrality provides a practical measuring method to answer these questions. Freeman (Freeman 1979) asserts that high centrality nodes should have three characteristics: they have more connections, can rapidly reach other nodes, and control the flow between other nodes (e.g., the center node in a star graph controls the flow between the other nodes). Based on these three characteristics, Freeman proposed three centrality measures: degree centrality, closeness centrality, and betweenness centrality. In addition, a number of other importance indicators were also proposed, e.g., radiality centrality (Valente and Foreman 1998) and stress centrality (Shimbel 1953). The centrality measures are described briefly as follows:

- Degree centrality: Degree centrality is defined as the number of direct connections between a node and other nodes. Its objective is to measure this node's interaction and development with other nodes. If a node has a higher number of direct connections with other nodes, then this node generally will have more influence in this network. This also means that this node is more active in this network. In a directed network, this can be divided into two types of degree centrality: in-degree centrality and out-degree centrality.
- Closeness centrality: Closeness centrality is a closeness that measures a node in relation to other nodes, and is calculated based on the shortest path between two nodes. It is defined as the reciprocal of the average length of the shortest paths from the node in question to other nodes. Thus, the shorter the average path is from the node to other nodes, the higher the closeness centrality would be. Closeness centrality can represent how long it takes for a message to be transmitted from a certain node to all the other nodes. The closeness centrality $C_c(s)$ of node s is evaluated as follows:

$$C_c(s) = 1 / \left(\frac{\sum_{i \neq s} \sigma_{si}}{n-1}\right) \quad (1)$$

where σ_{si} is the length of the shortest path between node s and node i; n is the total number of nodes.

- Betweenness centrality: In a network, if a node is located on the shortest path between many other nodes, this node can be considered as having an important position. Thus, this node can control the interaction of other nodes. The detailed calculation method will be described in Subsection 3.2. Betweenness centrality differs from closeness centrality in that closeness centrality does not consider the control ability, and only considers how many connections to other nodes a node has. Betweenness centrality can measure a node's level of control over others. The higher a node's control level is in a network means that it has a high betweeness centrality. This also means that more nodes need to pass this node to reach other nodes.
- Radiality centrality: This indicator is also used for evaluating importance based on the shortest path. The evaluation methods for radiality centrality and closeness centrality are similar, and can be considered a type of closeness centrality. However, radiality centrality also considers the longest path in this network. The calculated value is between 0 and 1. Radiality centrality is a distance indicator for the distance from one node to the center node. The lower the value is, the more distance the node in question is from the center. Conversely, nodes closer to the center have a higher radiality centrality. The radiality centrality $C_r(s)$ for node s is calculated as follows:

$$C_r(s) = (D - \frac{\sum_{i \neq s} \sigma_{si}}{n - 1} + 1) / D \tag{2}$$

  where D represents the network's longest path.
- Stress centrality: The concept of stress centrality is similar to that of betweenness centrality, i.e., both evaluating the node's control ability. Betweenness centrality considers whether there is a replacement shortest path, while stress centrality does not. Stress centrality is the sum of all the shortest paths that pass through the node in question. The more paths that pass through the node in question, the more paths it controls.

2.3 Centrality based visualization

Although relatively less attention has been paid on small-word networks, a lot of works has been done on centrality-based visualization of other types of networks. Jia *et al.* (Jia et al. 2008) used centrality to improve filtering and concentrated link layout drawing methods for scale-free networks. As a scale-free network has fewer clusters, their method retains links with high betweenness centrality in order to produce an effective scale-free network visualization system. Shi (Shi et al. 2009) proposed the HiMap system, which uses hierarchical grouping and summary clustering to visualize graphs. The HiMap method uses a new self-adaptive data loading technique to accurately control the visual density of each visualization, which involves two layout optimization algorithms. One of the optimization methods is implemented with the betweenness centrality to effectively avoid visual chaos previously common to social networks. Correa *et al.* (Correa et al. 2012) used centrality sensitivity as an important indicator to help simplify the network and retain the main structure, and to provide a mechanism for evaluating networks with high uncertainty as a network indicator. Crnovrsanin *et al.* (Crnovrsanin et al. 2014) used two analysis techniques (including centrality sensitivity analysis and modularity clustering analysis) and multiple visualization techniques (including edge bundling, centrality-based graph layout, radial layout, and parallel coordinates) to render social networks, and applied it in a real-life friendship network.

2.4 Planar graph

In graph theory, a planar graph is a graph that can be drawn on a plane without edge crossings. Because it does not have crossing links, it is very intuitive and easy for human vision to identify node relationships and structural patterns. This is also why this work chose to use planar graphs for drawing networks. Finding the maximal planar subgraph has already been proven to be a NP-hard problem. This work therefore used the heuristic algorithm proposed by Junger *et al.* (Junger et al. 1998) for this purpose. This algorithm finds the maximum planar subgraph by removing the least number of edges and is known to be one of the best heuristic algorithms for solving this type of problems. Previous studies have also proposed many different methods for drawing planar graphs, including non-directional drawing, right angle drawing (Fößmeier and Kaufmann 1996), mixed-mode drawing (Gutwenger and Mutzel 1997), and straight-line drawing (Aleardi et al. 2013; Kant 1996). The planar graph can be drawn as a node-link diagram with straight-line edges. For straight-line graph drawing, Kant *et al.* (Kant 1996) introduced a method to optimize the criteria of the required drawing area, the smallest angle, and the number of curves on a grid plane.

2.5 Force-directed method

The force-directed method simulates a graph as a physical system with spring-like attractive forces and repulsive forces like those of electrically charged particles (Eades 1984), to place nodes so that all the edge lengths are as equal as possible and reduce the number of edge crossings as much as possible. Generally, there are three types of repulsive forces in conventional force directed methods: node-node repulsion, node-edge repulsion, and edge-edge repulsion. The node-node repulsion model (Eades 1984) replaces nodes by charged particles, and edges by springs. Then, the algorithm changes positions of nodes to reach a status in

which the sum of the spring attraction and charged particle repulsion is zero, and this produces a balanced layout structure.

To prevent a node from getting too close to an edge, node-edge repulsion based methods was also proposed. Davidson *et al.* (Davidson and Harel 1996) applied the paradigm of simulated annealing in graph drawing and turned it into a combinatorial optimization problem, which sets a cost function for the distance between each node and edge. Penalties are given to the nodes and the edge if they are too close to each other. This method aims to find the optimal layout that meet multiple quality criteria. Bertault (Bertault 2000) proposed a node-edge repulsion based method that preserves edge crossings of the initial layout and ensures no new edge crossings being introduced. This method guarantees that the resulting drawing has no crossings if the original layout is planar. This approach is also adopted as part of our method. Lin *et al.* (Lin and Yen 2012) proposed an edge-edge repulsion method that can overcome zero angular resolution problems, therefore guaranteeing no overlapping of neighboring edges of a node.

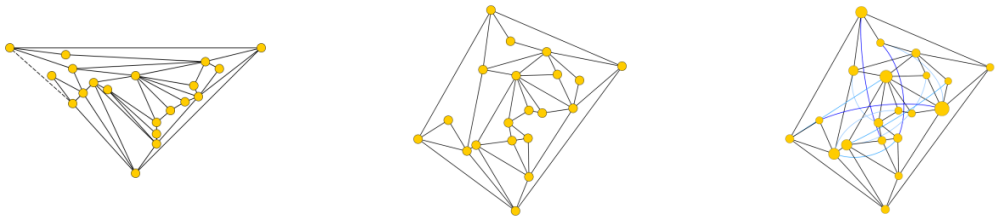
3. Methodology

Given the complexity of relationship patterns in a small-world network, it is often difficult to find out how clusters and entities interact with each other in its original data form. Thus, a method is needed to produce small-world network visualizations that can help network analysts to perform at least following tasks effectively: identifying groups, identifying center actors, and analyzing the role and responsibilities of those actors. To achieve these three main network analysis tasks, the proposed visualization method is a procedure that starts from a random drawing to the final graph drawing (as shown in Fig. 3(c)) in which users are able to rapidly retrieve embedded information and understand the underlying network data. The proposed algorithm is given in Algorithm 1. Each step of the algorithm is detailed in the following subsections.

Algorithm 1. Drawing a small-world graph

Input: A small-world network (V, E).
Output: A drawing for the network.

Step 1: Calculate the betweenness centrality of each node and link.
Step 2: Assign higher weights to links with low link betweenness centrality scores. Use the PQ-tree method (Junger et al. 1998) to find a minimum-cost (i.e., maximum-weight) planar subgraph (V, E_1). Note that the subgraph has the same number of nodes with the original network graph, but may have a smaller number of links.
Step 3: Without violating the graph planarity, use the algorithm in (Boyer and Myrvold 2004) to add to the subgraph as many links as possible according to the non-increasing ordering of link weights. Denote this new subgraph as (V, E_2). Note that $|E_2| \le |E_1| \le |E|$. Then, use the plane graph drawing algorithm (Kant 1996) to produce the planar graph drawing (e.g., Fig. 3(a)).
Step 4: Use the force-directed method based on node-link repulsion (Bertault 2000) to redraw the planar graph (V, E_2). Note that this drawing method still preserves the graph planarity (e.g., Fig. 3(b)).
Step 5: For each link in the set $E \setminus E_2$ in the non-decreasing ordering of link centrality, we draw this link as a curve to avoid from generating more crossings in the drawing. These added links are colored from dark to light to indicate the degree of link centrality, and the size of each node in V reflects its node centrality (e.g., Fig. 3(c)).



(a) Step 3 (b) Step 4 (c) Step 5

Fig. 3. Our method is used to draw a simple graph with 20 nodes and 60 links. (a) The drawing after Step 3, in which the added link is represented as a dashed line. (b) The drawing after Step 4, which draws the graph using the force-directed method based on node-link repulsion. (c) The drawing after Step 5, which adopts curves to add back the other links and uses different node sizes to represent the centrality degree of each node.

3.1 Evaluating betweenness centrality

Consider drawing a small-world network $G = (V, E)$ where V is the set of nodes, and E is the set of links. Let a link connected to two end nodes $x, y \in V$ be denoted as $e(x, y)$. Since this work considers the network to be non-directional, thus $e(x, y) = e(y, x)$. A path that starts with node x and finishes with node y, and passes through nodes from v_1 to v_n is denoted as $\langle x, e(x, v_1), v_1, e(v_1, v_2), v_2, \ldots, v_n, e(v_n, y), y\rangle$. Remind that the

shortest path from node s to node t denoted as σ_{st}. Note that $\sigma_{st} = \sigma_{ts}$, and $\sigma_{ss} = 1$. Also, let the shortest path from node s to node t that includes node u denoted as $\sigma_{st}(u)$. Then, the betweeness centrality of node u is evaluated as follows:

$$C_v(u) = \sum_{s \neq t \in V} \frac{\sigma_{st}(u)}{\sigma_{st}} \tag{3}$$

Let the shortest path from node s to node t that includes link k denoted as $\rho_{st}(k)$. Then, the betweenness centrality of link k is evaluated as follows:

$$C_e(k) = \sum_{s \neq t \in V} \frac{\rho_{st}(k)}{\sigma_{st}} \tag{4}$$

The time complexity of evaluating betweenness centrality on unweighted graph is $O(|V| \cdot |E|)$, while that on weighted graph is $O(|V| |E| + |V|^2 \log |V|)$ (Brandes 2001).

3.2 Identifying and drawing a planar subgraph

A link with high betweenness centrality means that many shortest paths pass through this link. That is, this link serves as a bridge between clusters. Conversely, links with low betweenness centrality are usually links within a cluster. Therefore, to maintain the linking structure within a cluster, this work uses low betweenness centrality links to draw the required planar graph, meaning that Step 2 of Algorithm 1 assigns higher weights to links with low link betweenness centrality scores. With the setting of weights, Step 2 of Algorithm 1 uses the PQ-tree algorithm of (Junger et al. 1998) to find a minimum-cost (maximum-weight) planar subgraph, which has the same number of nodes with the inputted graph, but has a smaller number of links. Next, without violating planar characteristics, Step 3 of Algorithm 1 applies the algorithm in (Boyer and Myrvold 2004) based on deep-first search to choose links and add the remaining links that are not in the planar subgraph back to the graph according to the non-increasing ordering of link weights (e.g., see dashed blue lines in Fig. 4). Note that the result obtained by the algorithm in (Boyer and Myrvold 2004) is a planar graph, not a plane graph drawing. Hence, we further use the plane drawing algorithm proposed by Kant (Kant 1996), which generates a plane drawing with straight-line links (e.g., Fig. 4).

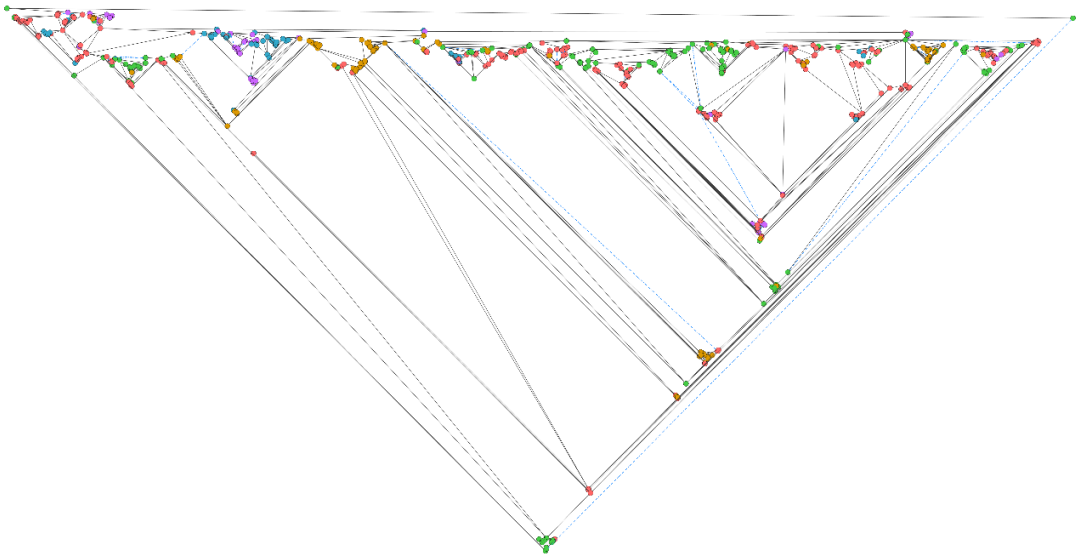


Fig. 4. The drawing after executing Step 3 of the proposed method on the graph in Fig. 1. The drawing has 732 links, in which the process of planarization has added 68.4% of the links of the original graph. In the drawing, the 36 links added from Step 2 to Step 3 are represented by blue dotted links.

3.3 Using the force-directed method based on node-edge repulsion

After creating a plane drawing of the planar subgraph, Step 4 of Algorithm 1 takes the plane drawing as the input, and feeds it into a force-directed method based on node-edge repulsion (Bertault 2000) to improve the plane drawing. This method takes three types of forces into consideration: node-node attraction for linked nodes, node-node repulsion for non-linked nodes, and node-edge repulsion for some conditions.

Let the position of node v denoted as $p(v)$, the distance between two nodes u and v as $d(u, v)$, and the edge length that we hope to obtain as δ. The attraction $F^a(u, v)$ and repulsion $F^r(u, v)$ between nodes u and v are respectively evaluated as follows:

$$F^a(u, v) = \frac{d(u,v)}{\delta}\big(p(v) - p(u)\big) \tag{5}$$

$$F^r(u, v) = \frac{-\delta^2}{d(u, v)}\big(p(v) - p(u)\big) \tag{6}$$

To calculate the repulsion $F^e(v, (a, b))$ between node v and edge (a, b), we consider a virtual node i_v, which is a vertical projection of node v on edge (a, b). This repulsion is set to only act on distance less than parameter γ between v and i_v. In this work, parameter γ is set to be four times the value of δ. Or, when $v = a$ or $v = b$, this repulsion is omitted. Hence, the repulsion $F^e(v, (a, b))$ between node v and edge (a, b) is evaluated as follows:

$$F^e(v, (a, b)) = \begin{cases} -\dfrac{(\gamma - d(u, v))^2}{\delta\, d(v, i_v)}(x(i_v) - x(v)), & \text{if } i_v \in (a,b),\ d(v,i_v) < \gamma, \\ & v \neq a, \text{ and } v \neq b; \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

Therefore, the total force applied on node v is as follows:

$$F(v) = \sum_{(u, v) \in E} F^a(u, v) + \sum_{u \in V} F^r(u, v) + \sum_{(a, b) \in E} F^e(v, (a, b)) - \sum_{u, w \in V, (v, w) \in E} F^e(u, (v, w)) \tag{8}$$

The time complexity of each iteration of the method is $O(|V|^2 + |V||E|)$. After several iterations, this algorithm can produce a more balanced planar drawing. For example, as seen in Fig. 5, the drawing produced at this step is a much better visualization. By looking at the drawing, we can easily detect the clusters in the graph and their distribution. Also since there is no overlapping of nodes and links as a result of the algorithm used, we can clearly identify paths from one node to another.

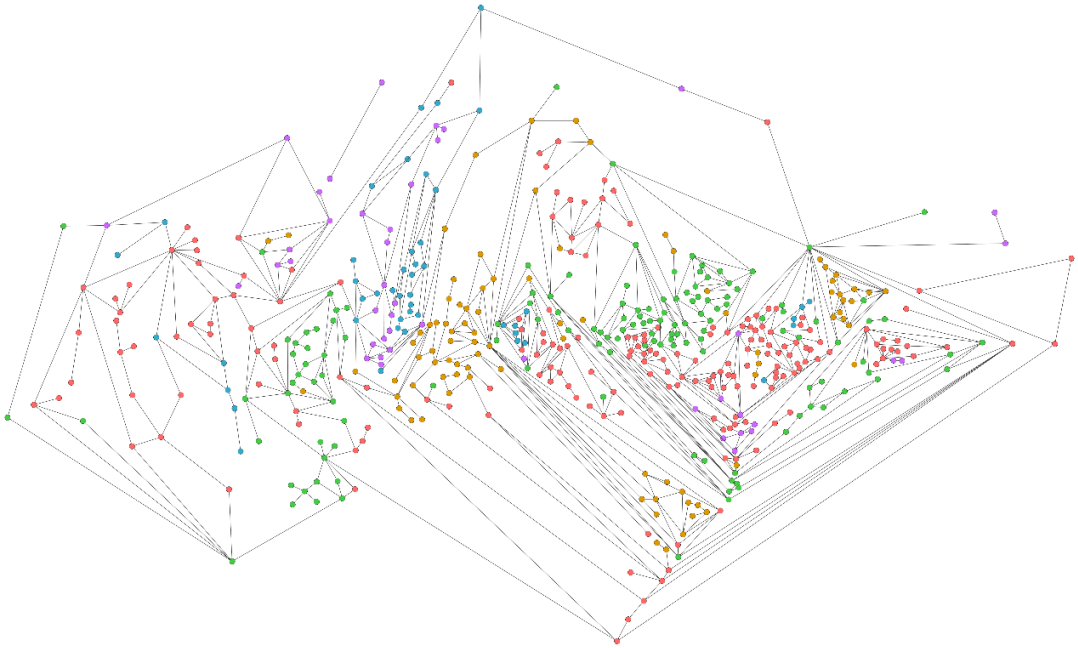


Fig. 5. The drawing after using the graph drawing method based on node-edge repulsion at Step 4.

3.4 Putting back the remaining links

After the plane graph is obtained at Step 4, Step 5 adds back the links that were deleted in Step 2 to produce the final drawing of the inputted network. We observe that those links that had high centrality scores and were left out at Step 2 were mostly bridges between clusters. This meant that the edge length of the deleted links was rather far. If those links were added back directly as straight lines, it could cause node and line overlapping and visual clutters. Therefore, as the previous work in (van Ham and Wattenberg 2008), the links are drawn back as curves. Different from (van Ham and Wattenberg 2008), this work emphasizes important links (i.e., bridge links between clusters), and hence adds back only those high centrality bridge links back. Furthermore, those links that are added back are colored according to the link centrality. In addition, the curves were overlaid on the plane drawing produced in the previous step. This way not only highlights important links, but also does not damage the drawing that we produced. Finally, node size is used to indicate

importance of nodes in terms of node betweenness centrality. That is, a larger node indicates that it has a higher centrality value. The final drawing is shown in Fig. 1(c).

4. Experiment analysis

Aside from comparing our resulting drawing with that of the work in (van Ham and Wattenberg 2008) that this work improves upon, this work also conducts experiments to quantitatively test the proposed method based on a series of graph and visualization parameters. The experimental instance is the Genealogy of Influence (Love 2007), a real-world network data. This instance is a small-world network with 515 nodes and 1070 links, describing how historically well-known intellectuals influence each other. A drawing example of this network produced by the proposed method is shown in Fig. 1(c). The nodes represent people, while the node color indicates profession, which includes artists (beige), philosophers (red), writers (green), scientists (blue), and mathematicians (purple). The links represent the effect that individuals have on each other, e.g., guiding others, building on other people's work or inheriting the thinking.

4.1 Basic parameter analysis

This subsection conducts some basic parameter analysis (Dong and Horvath 2007) on the structures of the original graph and the two graphs obtained by Step 3 and Step 5 of Algorithm 1. We calculate the following parameters, and the results are shown in Table 1:

- Average cluster coefficient: A network is called a small-world network if its average clustering coefficient is far greater than the average clustering coefficient of the random graph structure for the same collection of nodes, and if its average shortest path length is approximately the same as that in the random graph. In a non-directional graph, the cluster coefficient of node u is $2\, e_u / (k_u \cdot (k_u - 1))$, in which k_u is the number of neighbors that node u has, and e_u is the number of links that the neighbors have. The average cluster coefficient of a random graph with 515 nodes and 1070 links is approximately 0.0031. From Table 1, the graph structure from Step 3 has the highest cluster coefficient value (0.130671), because it was built mainly from in-cluster links. In Step 5, bridge links are added back, so this coefficient becomes lightly lower (0.122008), which is slightly higher than that of the initial graph.
- Network diameter: This is the path length of the longest path in the network and serves as the basis for the initial network size measurement. Since our subgraph in Step 3 is build based on low centrality in-cluster links and some more links added. On the other hand, more between-cluster bridge links are added in Step 5. This means that the graph of Step 3 has the least number of bridge links, thus having the largest network diameter value.
- Average degree, or connectivity: Average degree is one of the most basic parameters for assessing network complexity, and is given by the total number of links divided by the number of nodes. From Table 1, the graph obtained by the proposed method (i.e., the Step 5 graph) is one degree less than the initial graph for each node on average.
- Network density: This is a ratio of the number of current links to the number of possible links in a network if the network is a complete graph. Because the experimental data is a small-world graph, this indicator is not high for all three structures. Let n denote the number of nodes. Then, the network density is evaluated as follows:

$$Density = \frac{mean(Connectivity)}{n-1} \tag{9}$$

- Centralization: The closer the graph's topographical structure is to a star shape, the closer this parameter is to 1. The centralization value of the Step 5 graph (0.025258) is closer to that of the original graph (0.030954), compare to that of the Step 4 graph (0.021799). The parameter is evaluated as follows:

$$Centralization = \frac{n}{n-2}\left(\frac{max(Connectivity)}{n-1} - Density\right) \tag{10}$$

- Average shortest path length: We can see that the average shortest path length of the three structures is not long in general, which conforms to the small-world graph characteristic.
- Network heterogeneity: This is to assess the differences in the number of neighbors that each node has. The higher the value the greater the difference. Again, the network heterogeneity of the Step 5 graph is closer to that of the initial structure than that of the Step 3 graph. This parameter is evaluated as follows:

$$Heterogeneity = \frac{\sqrt{variance(Connectivity)}}{mean(Connectivity)} \tag{11}$$

As described by the aforementioned analysis, we can see that the graphs that we have drawn are already very close to the initial structure. In addition, we also analyze the distributions of the following parameters for the three structures, and the results are shown in Fig. 6, which is analyzed as follows: (1) Node degree: From

Fig. 6(a), the nodes with degree less than 10 in the initial graph are more likely to have their degree reduced in our drawing method. In particular, after the completion of our method, the number of nodes with degree from 1 to 3 increased the most. (2) Shortest path length: Fig. 6(b) shows the single peak distribution of all graphs. The original graph's curves are higher and the peak appeared at 5, while Step 3 curve is gentler and the peak appeared at 9. The peak for Step 5 curve appeared at 6 and was closer to the distribution of the original graph. (3) Cluster coefficient: This is the indicator used to measure the cluster relationship. The higher the connection level between node neighbors, the higher the cluster coefficient. As can be seen from Fig. 6(c), cluster coefficient distributions of Step 3 and Step 5 graphs are similar, but Step 5 is still relatively closer to the original graph. (4) Topological coefficient: This indicator is used to measure the level of each node sharing neighbors with other nodes. The higher this number is the more likely that the node is in the same cluster as its neighbors. As shown in Fig. 6(d), this distribution is displayed with the exponential trend line, and it can be seen that this indicator was not significantly different for the three graphs.

Table 1 Basic parameter analysis.

Parameter	Initial	Step 3	Step 5
Node count	515	515	515
Edge count	1070	732	792
Clustering coefficient	0.110529	0.130671	0.122008
Network diameter	14	20	16
Average degree	4.151456	2.838835	3.067961
Centralization	0.030954	0.021799	0.025258
Average shortest path length	5.084931	8.987677	6.399811
Network density	0.008077	0.005523	0.005969
Network heterogeneity	0.784044	0.640307	0.718529

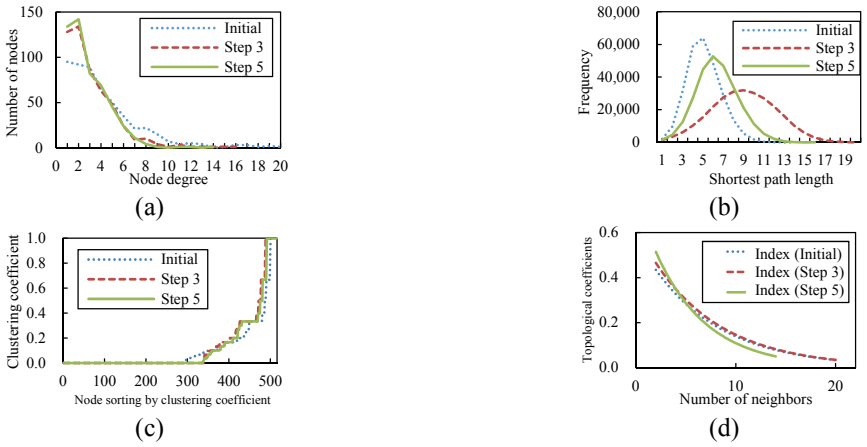


Fig. 6. (a) Node degree, (b) shortest path length, (c) cluster coefficient, and (d) topological coefficient.

4.2 Centrality analysis

We present the analysis for the three graph structures in regard to closeness centrality, radiality centrality, stress centrality, and betweenness centrality, and the results are shown in Fig. 7. (1) Closeness centrality (Fig. 7(a)): This indicator measures the average shortest path from one node to another node. It is the reciprocal of the average shortest path to other nodes. The higher this number, the shorter the average path. Generally, a node with more neighbors tend to have higher closeness centrality. The linear trend lines of the three graphs show that the trend of the Step 5 graph is close to that of the original graph, but the Step 3 graph is significantly different. (2) Radiality centrality (Fig. 7(b)): This indicator is also another method for evaluating centrality based on the shortest path. Its value range is between 0 and 1. It can be seen that the radiality centrality of the three structures is generally between 0.4 and 0.8. In Step 5, the important links were put back, so the Step 5 graph structure is more similar to the original graph than the Step 3 graph. (3) Stress centrality (Fig. 7(c)): Stress centrality is the sum of all the shortest paths that pass through a node. The more the shortest paths that pass through the node, the more paths that the node controls. Because the three graphs have different numbers of links, the stress centrality is normalized before making comparison. From Fig. 7(c),

although the climbing speed of the Step 3 and Step 5 graphs is not as high as the original graph, the general trend is the same. Not surprisingly, the value for Step 5 is still superior to that for Step 3. (4) Betweenness centrality (Fig. 7(d)): This indicator is the primary reference indicator in this work. The measurement of this centrality is similar to stress centrality. However, betweenness centrality considers the proportion of shortest paths passing through. For nodes with high stress centrality, if any two nodes have a high number of alternative shortest paths that do not include the node in question, then the node would have lower betweenness centrality.

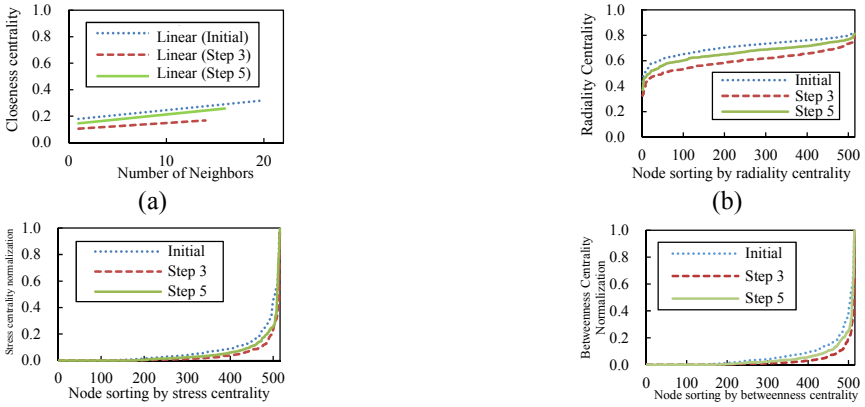


Fig. 7. (a) Closeness centrality, (b) radiality centrality, (c) stress centrality, and (d) betweenness centrality.

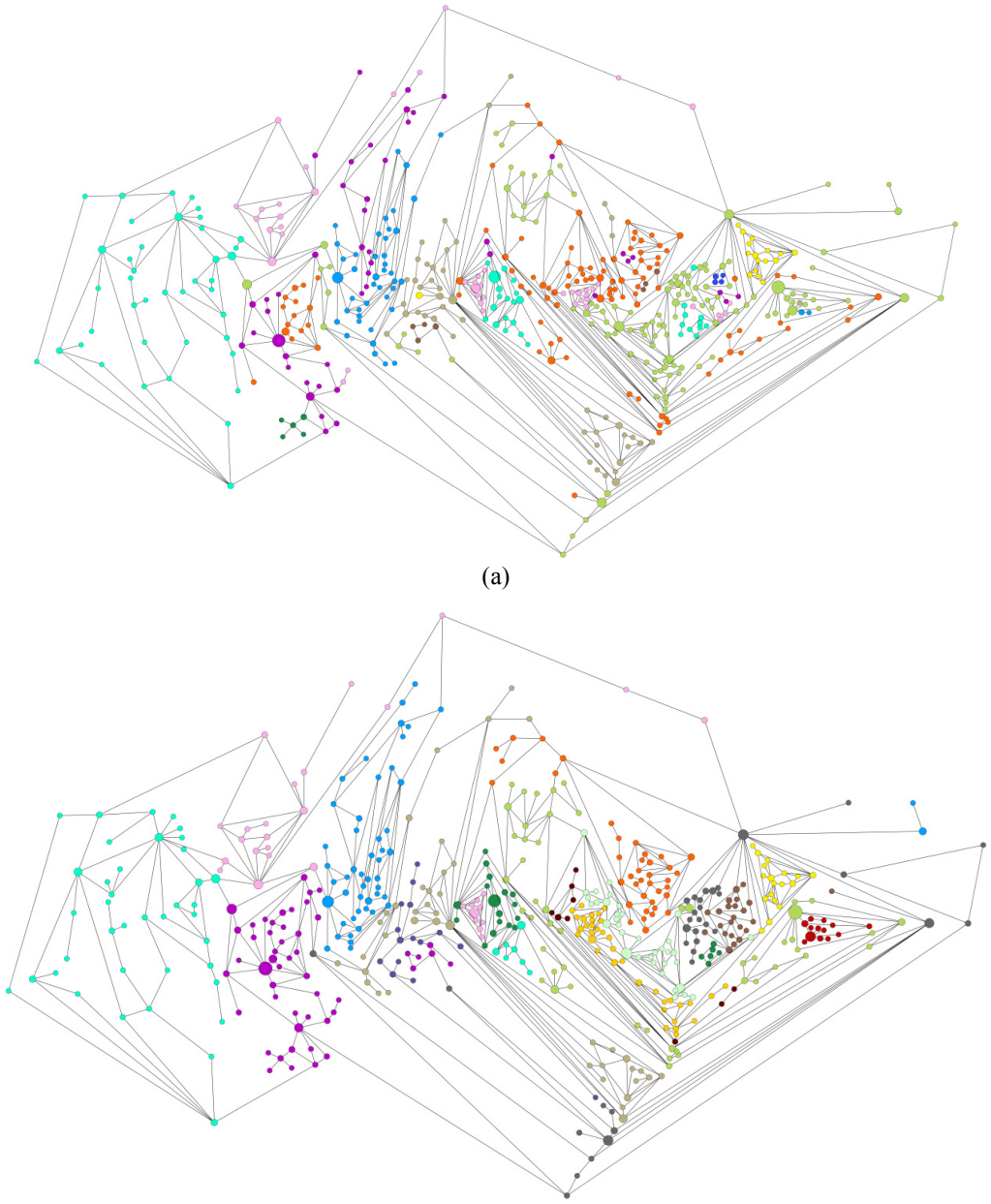
4.3 Clustering analysis

This subsection uses the clustering methods proposed by Girvan-Newman (Newman and Girvan 2004) and Wakita-Tsurumi (Wakita and Tsurumi 2007) to compare and analyze the three graph structures. The parameters that we use include the number of clusters, the number of nodes in the largest cluster, the number of nodes in the smallest cluster, the links within a cluster, and the number of bridge links between clusters. And the results are shown in Table 2.

Table 2 Clustering analysis of three structures.

Clustering method	Attribute	Initial	Step 3	Step 5
	Nodes	515	515	515
	Edges	1070	732	792
Girvan-Newman	Clusters	14	18	18
	Max cluster nodes	97	42	46
	Min cluster nodes	5	10	14
	Cluster edges	849	672	678
	Bridge edges	221	60	114
Wakita-Tsurumi	Clusters	21	24	25
	Max cluster nodes	65	39	43
	Min cluster nodes	9	6	11
	Cluster edges	732	660	657
	Bridge edges	308	72	135

From Table 2, the links added in Step 5 are almost all bridge links. This also verifies that as proposed in Subsection 3.2, links that are added back are indeed the ones with highest betweenness centrality values. In addition, the two clustering methods indicated both the Step 3 and Step 5 graphs have more groups than the original graph has. This is likely the result of selecting low betweenness centrality links in building our planar subgraph and maintaining the planarity of it during the process. The planarization can separate one cluster in the original graph into two clusters. However, this phenomenon does not seem to have a significant impact on the display of the clusters visually as shown in Fig. 8. The two pictures in Fig. 8 are generated using the clustering method of Girvan-Newman (Newman and Girvan 2004) on the original graph and the Step 5 graph. This method uses same colors to represent nodes in the same cluster. Fig. 8(a) shows the clustering of the original graph, while and Fig. 8(b) shows the clustering of the Step 5 graph.



(a)

(b)

Fig. 8. Comparison between (a) the clustering of the original structure and (b) the clustering of the Step 5 structure.

4.4 Visual analytics

We introduce phenomena or characteristics, or insights observed based on the final drawings generated by the proposed method. Fig. 9 shows a lot of visualization phenomena, e.g., connectivity structure and paths between nodes being easy to discern because the number of node and line overlaps and clutters decrease. Further, the positions of nodes in Fig. 9 from left to right reflect the years of the people from past to present. This is also the real world situation as people are related to each other in the same years. More specifically, the left-most side of the pictures shows the early Greek scholars, including writers, philosophers, and scientists.

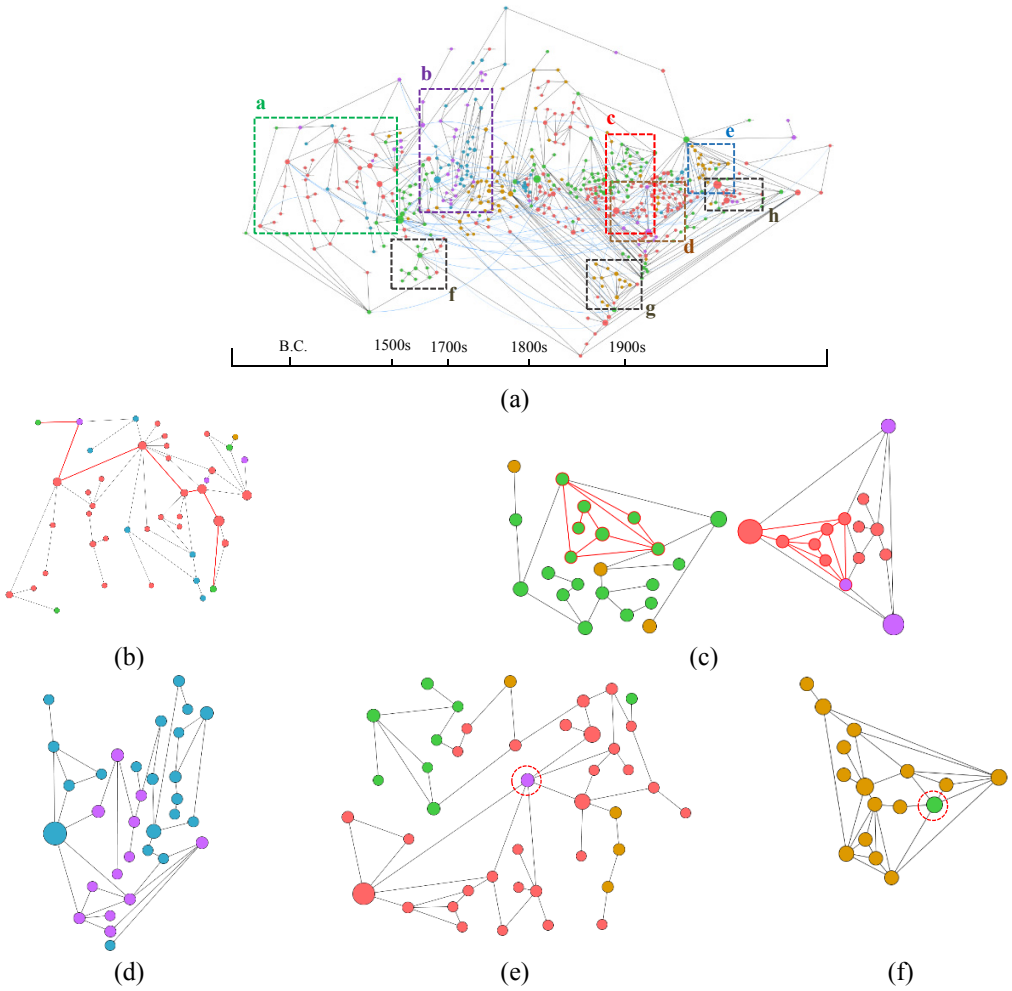


Fig. 9. (a) Final drawing using the proposed method, (b) enlarged drawing of region a, (c) enlarged drawing of region c, (d) enlarged drawing of region b, (e) enlarged drawing of region d, and (f) enlarged drawing of region e.

In regard to clustering relationships, the play writers are on the bottom left side of the middle area. We can see that Shakespeare had the most links. This means that he was influenced by a lot of people, and that he influenced a lot of people as well as shown in area f of Fig. 9. On the right side of the middle area, the figure shows a high connectivity between mathematicians and philosophers and logicians. The figure shows that the artists on the right side are not connected to the artists in the center. These artists form a cluster of composers, as shown in area e of Fig. 9. The artist cluster in the center is formed mostly by painters, in which the node with the most number of links is Picasso. The cluster in area g of Fig. 9 is formed by Dutch and French painters from the end of the 19th century to the 20th century. On the far right side is a small cluster formed by philosophers. We can see that they were mostly from the phenomenological school, as shown in area h of Fig. 9. Phenomenology is one of the most important philosophy schools in the 20th century. This philosophy was proposed by the German philosopher Husserl.

Other than basic cluster relationships, we can see certain special phenomena such as hierarchy structure, as seen in area c of Fig. 9. We can see that closer connections exist within certain clusters. Also, two different types of professions are included in the same cluster. For example, area b of Fig. 9 shows a close relationship among post 18th century scientists and mathematicians. In addition, from our visualization, we can also more special nodes, such as Descartes and Goethe. They were linked to more people from different professions. Collection of their background information reveals that this is because these two people worked in a number

of different professions. Another special node is Wittgenstein, who is the area d of Fig. 9. In the raw data, he is classified as a mathematician, but he is almost exclusively linked to philosophers, which is very interesting. A further investigation revealed that his research was focused mainly on mathematical philosophy. Similarly, another node in area e of Fig. 9 is Heine. Heine's profession is a poet, but almost all of his connections are composers. We later find that this is because Heine's poems were liked by many people, and many composers used his works to compose songs, thereby, creating the links from composers. Finally, we were also curious why links existed between different groups. Thus, we look at and analyze writers and philosophers near the center of the visualization. We find that Shelley is Godwin's daughter, that Stein is a student of Santayana and James, and that the works of Borges and Proust were influenced by Carlyle. These relationships cause the writers and philosophers to become neighbors.

As demonstrated from the aforementioned analysis, we discover much information from the visualization that our method produced, which was otherwise difficult, or not able to find based on the data in its original format or visualizations from previous studies. We were also able to use the information found to gain more insights about the original graph and understand the relationships among these people.

5. Application

This work applies our drawing procedure on the semiconductor industry. The research data mainly came from a well-known semiconductor company in Taiwan. The semiconductor wafer production process often includes hundreds of processing stages. During processing, the processing information is automatically recorded as the basis for the next analysis. The following important characteristics during each processing stage for each wafer are recorded: processing time (date), factory area (fab), machine (tool), parameter allocation (recipe), carrier (carrier), and reaction chamber (chamber). More detailed descriptions of these data are include in Table 3. Because the semiconductor production process is very complex, it is difficult to find useful information from so much data. We hope to build a network that includes wafer information and use our visualization method to actually help analyze data of the semiconductor wafer manufacturing visually.

Table 3 Key features of manufacturing semiconductor wafer.

Feature	Content
Stage	Stage ID of the manufacturing process.
Date	The date of initializing the manufacturing process
Fab	The fabrication plant ID
Tool	Machine (tool) ID adopted at the stage
Recipe	The recipe of setting parameters
Chamber	Chamber ID of the machine adopted at the stage
Carrier	ID of the carrier that enters the chamber

In the network, each node represents a wafer. For the links between the wafers, we used the aforementioned wafer processing characteristics to calculate the processing similarity between any two wafers. We selected 1600 links between two nodes with the highest similarities to build a graph with 467 nodes and 1600 links. Our calculation indicated that the clustering coefficient of this graph is 0.336, which is greater than the cluster coefficient of a random graph (0.014). Also, the average path length is 9.529, which conforms to small-world graph characteristics. Subsequently, we used our proposed drawing method to visualize this graph. The number of the links in the final drawing was reduced to 1047, and the final drawing is shown in Fig. 10.

We used the Girvan-Newman clustering method (Newman and Girvan 2004) to cluster this graph and used different colors to indicate which cluster each node belonged to. The result revealed that some clusters had higher density, some clusters had lower density, while some were made of some dispersed nodes. The dense clusters are shown in a red frame in Fig. 10. Because the processing similarity of any two wafers in the cluster is highly similar, the production time can be sped up and the production cost can be reduced. The low-density clusters and clusters with dispersed nodes are shown in green frames in Fig. 10. Because these types of wafers have lower similarity with other wafers or wafer processes, it may mean that these wafers have lower processing performance compared to other wafers and require improvement. Therefore, we can use this graph to improve the wafer production procedure, find abnormal wafers and determine the cause. Compared to filling rich wafer data information in numerous forms, our visualization procedure and analytic approach can rapidly and effectively help users intuitively obtain information.

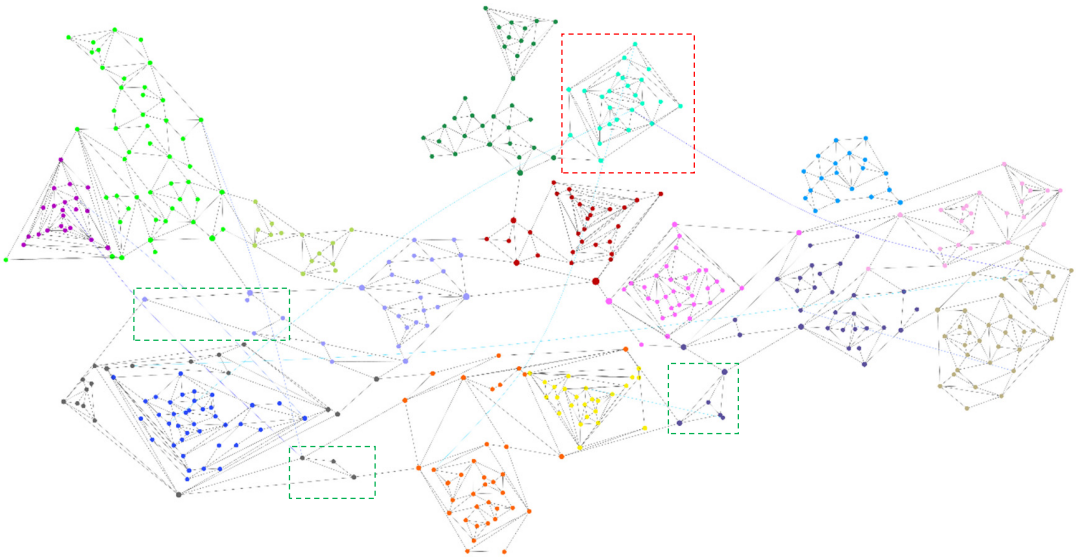


Fig. 10. The drawing of the graph for key features of manufacturing semiconductor wafer.

6. Conclusion

Different from conventional network visualization algorithms, this work provides an improved visualization method for small-world networks. In this method, we first use betweenness centrality scores to choose links and produce the initial weighted planar subgraph. The new force-directed method is then used to re-layout the graph. Finally, we placed back high centrality links to maintain the completeness of the network structure. In previous studies, very few of them have utilized planar graph drawing algorithms and force-directed methods based on node-link repulsion on small-world networks. Therefore, we provided a new approach for visualization and visual analytics of this type of data. For the experiments we used genealogy of influence (Love 2007) as the data example. We conducted a detailed basic parameter and centrality analysis on the structures of the original graph, the planar subgraph drawn with our method and the final graph. Notably, we conducted a clustering analysis using two methods to compare the clustering differences between the structures of the graphs. Based on the final visualization produced by our method, the analysis we conducted not only found the same information as that in (van Ham and Wattenberg 2008), but also successfully revealed more useful information. This gave us a better understanding of the relationships between notable intellectuals in history, and discovered some previously undiscovered information.

Finally, for future research directions, future studies may attempt to use different measurements or methods to produce the initial planar subgraphs. For example, a different centrality measure, rather than betweenness centrality, can be used to choose links for the planar graph. Further, different force-directed methods, edge binding, or other visualization techniques may also be used for visualization of the final graph to see if they could produce better results. In addition, we also hope that our method can be applied to small-world networks in different fields to explore and retrieve interesting and useful information.

Acknowledgements

The authors thank the anonymous referees for comments that improved the content as well as the presentation of this paper. This work has been supported in part by Ministry of Science and Technology, Taiwan, under Grants MOST 105-2628-E-009-002-MY3 and MOST 106-2221-E-009-101-MY3, and the Ministry of Education, Taiwan, under the Higher Education Sprout Project.

References

Abbasi A, Hossain L, Leydesdorff L (2012) Betweenness centrality as a driver of preferential attachment in the evolution of research collaboration networks. Journal of Informetrics 6:403-412

Aleardi LC, Devillers O, Fusy É (2013) Canonical ordering for triangulations on the cylinder, with applications to periodic straight-line drawings. In: Proc. of Graph Drawing, 2013. Springer, pp 376-387

Archambault D, Munzner T, Auber D (2011) Tugging graphs faster: Efficiently modifying path-preserving hierarchies for browsing paths. IEEE Transactions on Visualization and Computer Graphics

17:276-289
Bassett D, Bullmore E (2017) Small-World Brain Networks Revisited. Neuroscientist 23:499-516
Bender-deMoll S, McFarland DA (2006) The art and science of dynamic network visualization. Journal of Social Structure 7:1-38
Bertault F (2000) A force-directed algorithm that preserves edge-crossing properties. Information Processing Letters 74:7-13
Bertin J (1983) Semiology of graphics: diagrams, networks, maps. Esri Press.
Bhandari A, Gupta A, Das D (2017) Betweenness centrality updation and community detection in streaming graphs using incremental algorithm. In: Proc. of the 6th International Conference on Software and Computer Applications, 2017. ACM Press, pp. 159-164.
Bi C, Fu B, Chen J, Zhao Y, Yang L, Duan Y, Shi Y (2018) Machine learning based fast multi-layer liquefaction disaster assessment. World Wide Web doi:10.1007/s11280-018-0632-8
Bi CK, Yang L, Duan YL, Shi Y (2019) A survey on visualization of tensor field. Journal of Visualization 22:641-660
Bonchi F, Morales GDF, Riondato M (2016) Centrality Measures on Big Graphs: Exact, Approximated, and Distributed Algorithms. In: Proc. of the 25th International Conference Companion on World Wide Web, 2016. ACM Press, pp. 1017-1020
Borgatti SP, Mehra A, Brass DJ, Labianca G (2009) Network analysis in the social sciences. Science 323:892-895
Boyer JM, Myrvold WJ (2004) On the cutting edge: Simplified $O(n)$ planarity by edge addition. Journal of Graph Algorithms and Applications 8:241-273
Brandes U (2001) A faster algorithm for betweenness centrality. Journal of Mathematical Sociology 25:163-177
Brandes U, Indlekofer N, Mader M (2012) Visualization methods for longitudinal social networks and stochastic actor-oriented modeling. Social Networks 34:291-308
Burger M, Zelazo D, Allgower F (2013) Hierarchical clustering of dynamical networks using a saddle-point analysis. IEEE Transactions on Automatic Control 58:113-124
Correa C, Crnovrsanin T, Ma K-L (2012) Visual reasoning about social networks using centrality sensitivity. IEEE Transactions on Visualization and Computer Graphics 18:106-120
Crnovrsanin T, Muelder CW, Faris R, Felmlee D, Ma K-L (2014) Visualization techniques for categorical analysis of social networks with multiple edge sets. Social Networks 37:56-64
Das S, Lee D, Choi W, Doppa JR, Pande PP, Chakrabarty K (2017) VFI-based power management to enhance the lifetime of high-performance 3D NoCs. ACM Transactions on Design Automation of Electronic Systems 23:1-26 doi:10.1145/3092843
Davidson R, Harel D (1996) Drawing graphs nicely using simulated annealing. ACM Transactions on Graphics 15:301-331
Dong J, Horvath S (2007) Understanding network concepts in modules. BMC systems biology 1:24
Eades P (1984) A heuristics for graph drawing. Congressus numerantium 42:146-160
Fößmeier U, Kaufmann M (1996) Drawing high degree graphs with low bend numbers. In: Proc. of Graph Drawing, 1996. Springer, pp 254-266
Freeman LC (1979) Centrality in social networks conceptual clarification. Social networks 1:215-239
Gómez D, Figueira JR, Eusébio A (2013) Modeling centrality measures in social network analysis using bi-criteria network flow optimization problems. European Journal of Operational Research 226:354-365
Gibson H, Vickers P (2016) Using adjacency matrices to lay out larger small-world networks. Applied Soft Computing 42:80-92
Gretarsson B, O'Donovan J, Bostandjiev S, Hall C, Höllerer T (2010) Smallworlds: Visualizing social recommendations. Computer Graphics Forum 29:833-842
Guan C, Yuen KKF (2013) Towards a hybrid approach of primitive cognitive network process and k-means clustering for social network analysis. In: Proc. of IEEE International Conference on Internet of Things (iThings/CPSCom), 2013. IEEE Press, pp 1267-1271
Gutwenger C, Mutzel P (1997) Grid embedding of biconnected planar graphs. Extended Abstract, Max-Planck-Institut f ur Informatik, Saarbr ucken, Germany
Huang W, Eades P, Hong S-H, Lin C-C (2013) Improving multiple aesthetics produces better graph drawings. Journal of Visual Languages & Computing 24:262-272
Huang W, Huang ML, Lin C-C (2016) Evaluating overall quality of graph visualizations based on aesthetics

aggregation. Information Sciences 330:444-454
Jia Y, Garland M, Hart JC (2011) Social network clustering and visualization using hierarchical edge bundles. Computer Graphics Forum 30:2314-2327
Jia Y, Hoberock J, Garland M, Hart JC (2008) On the visualization of social and other scale-free networks. IEEE Transactions on Visualization and Computer Graphics 14:1285-1292
Junger M, Leipert S, Mutzel P (1998) A note on computing a maximal planar subgraph using PQ-trees. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems 17:609-612
Kant G (1996) Drawing planar graphs using the canonical ordering. Algorithmica 16:4-32
Lee T, Kim C (2014) Statistical comparison of fault detection models for semiconductor manufacturing processes. IEEE Transactions on Semiconductor Manufacturing 28:80-91
Lin C-C, Huang W, Liu W-Y, Chen W-L (2018) Evaluating aesthetics for user-sketched layouts of symmetric graphs. Journal of Visual Languages & Computing 48:123-133
Lin C-C, Yen H-C (2012) A new force-directed graph drawing method based on edge–edge repulsion. Journal of Visual Languages & Computing 23:29-42
Love M (2007) Genealogy of Influence. [Online]. Available: http://mikelove.nfshost.com/genealogy/
Newman ME, Girvan M (2004) Finding and evaluating community structure in networks. Physical review E 69:026113
Opsahl T, Agneessens F, Skvoretz J (2010) Node centrality in weighted networks: Generalizing degree and shortest paths. Social Networks 32:245-251
Shi L et al. (2009) HiMap: Adaptive visualization of large-scale online social networks. In: Proc. of 2009 IEEE Pacific Visualization Symposium, 2009. IEEE Press, pp 41-48
Shimbel A (1953) Structural parameters of communication networks. The bulletin of mathematical biophysics 15:501–507
Sohn K, Kim D (2010) Zonal centrality measures and the neighborhood effect. Transportation Research Part A: Policy and Practice 44:733-743
Valente TW, Foreman RK (1998) Integration and radiality: measuring the extent of an individual's connectedness and reachability in a network. Social Networks 20:89-105
van Ham F, Wattenberg M (2008) Centrality based visualization of small world graphs. Computer Graphics Forum 27:975-982
Wakita K, Tsurumi T (2007) Finding community structure in mega-scale social networks:[extended abstract]. In: Proc. of the 16th International Conference on World Wide Web, 2007. ACM Press, pp 1275-1276
Watts DJ, Strogatz SH (1998) Collective dynamics of ‘small-world’networks. Nature 393:440-442
Wong PC, Foote H, Mackey P, Chin G, Huang Z, Thomas J (2012) A space-filling visualization technique for multivariate small-world graphs. IEEE Transactions on Visualization and Computer Graphics 18:797-809
Yang Z, Chen W (2018) A game theoretic model for the formation of navigable small-world networks—the tradeoff between distance and reciprocity. ACM Transactions on Internet Technology 18:1-38
Zuo X-N, Ehmke R, Mennes M, Imperati D, Castellanos FX, Sporns O, Milham MP (2012) Network centrality in the human functional connectome. Cerebral Cortex 22:1862-1875