

Algorithms for Visualizing Large Networks

Yifan Hu
AT&T Labs – Research
180 Park Ave.
Florham Park, NJ 07932, USA
yifanhu@research.att.com

September 16, 2011

1 Introduction

Graphs are often used to encapsulate relationship between objects. Graph drawing enables visualization of such relationships. The usefulness of this visual representation is dependent on whether the drawing is aesthetic. While there are no strict criteria for aesthetics of a drawing, it is generally agreed, for example, that such a drawing has minimal edge crossing, with vertices evenly distributed in the space, connected vertices close to each other, and symmetry that may exist in the graph preserved.

One of the earliest example of graph drawing is probably the Mill (also known as Morris) game boards, found in the 13th century *The Book of Games*, produced under the direction of Alfonso X (1221-1284), King of Castile (Spain). Other examples of early graph drawing are drawings of family trees, and trees of virtues and vices.

In 1735, Euler presented and later published his famous paper [13] on the Seven Bridges of Königsberg. The problem was to find a walk through the city of Königsberg, which include two islands connected to the mainland by seven bridges (Figure 1). The walk must cross each bridge once and only once. Euler proved that such a walk is not possible. Euler's paper laid the foundation of graph theory. Interestingly, however, Euler did not provide a drawing of the graph representation of the problem itself in the paper.

In the 18th century, the Irish physicist, astronomer, and mathematician, William Hamilton, invented the Icosian game. To play the game, one has to visit every node of a dodecahedral graph once, and back to the starting node. Every edge can be passed at most once. In 1857 he sold the game to a London game dealer for 25 Pounds (about 3000 US dollars in today's money after taking inflation into account). The dealer made two versions; unfortunately, neither sold well, probably because the game was so simple that even a child could succeed. Around the same time, British mathematician Arthur Cayley wrote a paper on the enumeration of tree diagrams. In the paper he pioneered the

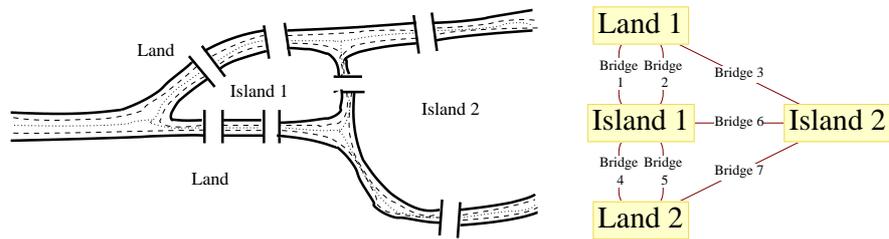


Figure 1: An illustration of the seven bridges of Königsberg (left) and the graph representation (right).

concept of a tree, which is now fundamental to computer sciences. The paper also contains drawings of trees.

Between the 18th and 19th centuries, graph drawing also appeared in areas outside of mathematics. For example, in crystallography and chemistry, graph drawings were used to illustrate molecular structures. In 1878, the English mathematician Sylvester introduced the concept of a *graph* for the first time.

From the mid-20th century, with the demands from the military, transportation and communication, and science and technology, a great number of graph theory problems emerged. Often it is helpful to be able to visualize a graph. However drawing a complex graph by hand is time consuming, if not impossible, for large graphs. Therefore automatic generation of graph drawings became of interest, and were facilitated by the ever increasing computing power.

In 1963, Tutte [55] proposed an algorithm to draw planar graphs by fixing nodes on a face and placing the rest of the nodes at the barycenters of their neighbors. In 1981, Sugiyama et al. [52] proposed a method for the hierarchical drawing of directed graphs. Eades [12] proposed a force-directed algorithm in 1984, and Kamada and Kawai [37] the spring embedding algorithm in 1989. These algorithms were further improved [20, 15], and some of them applied to large graphs [5, 24, 34, 49, 54, 57] later.

Depending on the applications and properties of the graph to be visualized, there are many styles of graph drawing. For example, for planar graphs, a planar embedding is more appropriate. For general graphs, there are also two broad styles of drawing: straight line drawing, where each edge is represented by a straight line, and orthogonal drawing, where each edge is represented by lines segments that are either horizontal or vertical. Finally, for directed graphs where it is important to show the direction of the edges, a hierarchical drawing style can be used. Figure 2 shows a graph drawn using the hierarchical style (left) and the undirected straight-edged style (right).

In this paper we look at algorithms for visualizing large networks. The term large is relative to the computing power and memory available. In this paper, by large, we mean graphs of more than a few thousand vertices. Such large graphs occur in areas such as Internet mapping, social networks, biological pathways and genealogy. Large networks bring unique issues. For example,

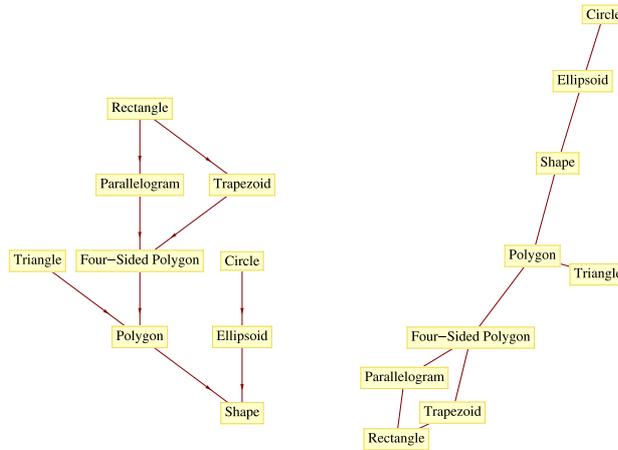


Figure 2: A graph drawn in hierarchical (left) and undirected (right) styles.

the complexity of the algorithm becomes very important. The ability of the algorithm in escaping from local minimum and achieving a globally optimal drawing is also vital. We present some of the algorithms suitable for large graphs. We shall limit our attention to straight edge drawing of undirected graphs, mostly because scalable algorithms for these have been more intensely investigated. Our emphasis is on algorithms, with a tilt toward issues relating to combinatorial computing. We do not attempt to give a comprehensive review of the literature and history, nor of software and visualization systems; for these the reader is referred to [3, 39].

We note that not all graph can be aesthetically embedded into two or three-dimensional space. For example, many Internet graphs are of “small-world” nature [60]. Such graphs are difficult to layout in Euclidean space, and often require an interactive system to comprehend and explore. A number of techniques can be brought to bear on such graphs, including visual abstraction [50, 18], fisheye-like distortion and hyperbolic layout [18, 31, 42, 46, 45]. These topics are not within the scope of this chapter, but are nevertheless very important parts of large graph visualization.

2 Algorithms for Drawing Large Graphs

We use $G = \{V, E\}$ to denote an undirected graph, with V the set of vertices and E the set of edges. Denote by $|V|$ and $|E|$ the number of vertices and edges, respectively. If vertices i and j form an edge, we denote that $i \leftrightarrow j$, and call i and j neighboring vertices. We denote by x_i the current coordinates of vertex i in d -dimensional Euclidean space. Typically $d = 2$ or 3 .

The aim of graph drawing is to find x_i for all $i \in V$ so that the resulting

drawing gives a good visual representation of the connectivity information between vertices. One way to solve this problem is to turn the graph drawing problem into that of finding a minimal energy configuration of a physical system. Within this framework, two popular methods, the spring-electrical model [12, 15], and the stress model [37], are the most well-known.

We first discuss the spring-electrical model, and the multilevel approach and force approximation techniques that make this model suitable for large graphs. We note that the multilevel approach is not limited to the spring-electrical model, but for convenience we introduce it in the context of this model. We then discuss the stress model and classical MDS (strain model), and look at attempts to apply these models to large graphs. Finally we present high-dimensional embeddings and Hall’s algorithm for large graph layout.

Within the graph drawing literature, the name “force-directed algorithm” has often been used with the both spring-electrical and stress models. Strictly speaking, the name “force-directed algorithm” is appropriate when we talk about a procedure that involves calculating forces exerted on vertices, moving them along the direction of the forces, and repeating until the system comes to an equilibrium state. Therefore we shall only use the term “force-directed algorithm” when referring to this iterative procedure.

2.1 Spring-electrical Model

The spring-electrical model [12] represents the graph drawing problem by a system of electrically charged vertices attracted to each other by springs; vertices also repel each other via electrical forces. Here, following [15], the attractive spring force exerted on vertex i from its neighbor j is proportional to the squared distance between these two vertices,

$$F_a(i, j) = -\frac{\|x_i - x_j\|^2}{K} \frac{x_i - x_j}{\|x_i - x_j\|}, \quad i \leftrightarrow j,$$

where K is a parameter related to the nominal edge length of the final layout. The repulsive electrical force exerted on vertex i from any vertex j is inversely proportional to the distance between these two vertices,

$$F_r(i, j) = \frac{K^2}{\|x_i - x_j\|} \frac{x_i - x_j}{\|x_i - x_j\|}, \quad i \neq j.$$

The energy of this physical system [47] is

$$E(x) = \sum_{i \leftrightarrow j} \|x_i - x_j\|^3 / (3K) - \sum_{i \neq j} K^2 \ln(\|x_i - x_j\|),$$

with its derivatives a combination of the attractive and repulsive forces. We note that there are other variations of this force model. See, e.g., [4, 12].

The spring-electrical model can be solved with the aforementioned force-directed algorithm by starting from a random layout, calculating the combined attractive and repulsive forces on each vertex, and moving the vertices along

the direction of the force for a certain step length. This process is repeated, with the step length decreasing every iteration, until the layout stabilizes. The following algorithm starts with a random, or user supplied, initial layout x .

Algorithm 1: Force-directed algorithm

- ForceDirectedAlgorithm(G, x, tol) {
 - $converged = FALSE$;
 - $step =$ initial step length;
 - while ($converged$ equals $FALSE$) {
 - * $x^0 = x$;
 - * for $i \in V$ {
 - $f = 0$;
 - for ($j \leftrightarrow i$) $f := f + F_a(i, j)$;
 - for ($j \neq i, j \in V$) $f := f + F_r(i, j)$;
 - $x_i := x_i + step * (f / \|f\|)$;
 - * }
 - * $step := update_steplength(step, x, x^0)$;
 - * if ($\|x - x^0\| < tol * K$) $converged = TRUE$;
 - }
 - return x ;
- }

This procedure can be enhanced by an adaptive step length updating scheme [7, 34], and usually works well for small graphs.

For large graphs, this simple iterative procedure is not sufficient to overcome the many local minima that often exist in the space of all possible layouts. Instead, a multilevel approach has to be used (Section 2.1.2). Furthermore, a nested space partitioning data structure is needed to approximate the all-to-all electrical force so as to reduce the quadratic complexity to $O(|V| \log |V| + |E|)$ (Section 2.1.1). Combining these two powerful techniques results in efficient implementations of the spring-electrical model [24, 34] that are capable of handling graphs of millions of vertices and edges [33].

2.1.1 Fast force approximation

Each iteration of the force-directed algorithm (Algorithm 1) involves two loops. The outer loop iterates over each vertex. Of the two inner loops, the latter involves calculation of all-to-all repulsive forces, and $O(|V|)$ force calculations are needed for every vertex. Thus the overall complexity is $O(|V|^2)$.

The repulsive force calculation resembles the n -body problem in physics, which is well-studied. One of the widely used techniques to approximate the repulsive forces in $O(n \log n)$ time with good accuracy, but without ignoring

long range forces, is to treat groups of far away vertices as supernodes, using a suitable data structure [2]. This idea was adopted by Tunkelang [54] and Quigley [49]. They both used an quadtree (2D) or octree (3D) data structure.

For simplicity, hereafter we use the term quadtree exclusively, which should be understood as octree for 3D. A quadtree data structure is constructed by first forming a square that encloses all vertices. This is the level 0 square. This square is subdivided into 4 squares if it contains more than 1 vertex, and forms the level 1 squares. This process is repeated until each square contains no more than 1 vertex. Figure 3 (left) shows a quadtree on the `jagmesh1` graph.

The quadtree forms a recursive grouping of vertices, and can be used to efficiently approximate the repulsive force in the spring-electrical model. The idea is that in calculating the repulsive force on a vertex i , if a group of vertices, S , lies in a square that is sufficiently “far” from i , the whole group can be treated as a supernode. Otherwise we traverse down the hierarchy and examine the four sibling squares. The supernode is assumed to be situated at the center of gravity of the group, $x_S = (\sum_{j \in S} x_j)/|S|$. The repulsive force on vertex i from this supernode is

$$f_r(i, S) = \frac{K^2 |S|}{\|x_i - x_S\|} \frac{x_i - x_S}{\|x_i - x_S\|}.$$

It remains to define what “far” means. Following [49, 54], the supernode S is far away from vertex i , if the width d_S of the square that contains the supernode is small, compared with the distance between the supernode and the vertex i ,

$$\frac{d_S}{\|x_i - x_S\|} \leq \theta. \quad (1)$$

This inequality is called the Barnes-Hut opening criterion, and was originally formulated by Barnes and Hut [2]. Here $\theta \geq 0$ is a parameter. The smaller the value of θ , the more accurate the approximation to the repulsive force, and the larger the number of force calculations. A typical value that works well in practice is $\theta = 1.2$.

The quadtree data structure allows efficient identification of all the supernodes that satisfy (1). The process starts from the level 0 square. Each square is checked, and recursively opened, until the inequality (1) is satisfied. Figure 3 (right) shows all the supernodes (the squares) and the vertices these supernodes consist of, with reference to vertex i located at the top-middle part of the graph. In this case there are 936 vertices, and 32 supernodes.

Under a reasonable assumption [2, 48] of the distribution of vertices, it can be proved that building the quadtree takes a time complexity of $O(|V| \log |V|)$. Finding all the supernodes with reference to a vertex i can be done in a time complexity $O(\log |V|)$. Overall, using a quadtree structure to approximate the repulsive force, the complexity for each iteration of the force-directed Algorithm 1 is reduced from $O(|V|^2)$ to $O(|V| \log |V|)$. This force approximation scheme can

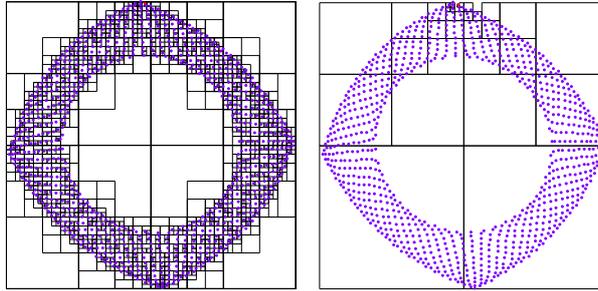


Figure 3: An illustration of the quadtree data structure. Left: the overall quadtree. Right: supernodes with reference to a vertex at the top middle part of the graph, with $\theta = 1$.

be further improved by considering force approximation at supernode-supernode level instead of vertex-supernode level [8].

A force approximation algorithm with the same $O(|V| \log |V|)$ complexity but that is independent of the distribution of vertices is the multipole method [22], Hachul and Jünger applied this force approximation to graph drawing [24].

2.1.2 Multilevel approach

While fast approximation of long range forces resolves the quadratic complexity of the force-directed algorithm for the spring-electrical model, it does not change the fact that the algorithm repositions one vertex at a time, without a “global view” of the layout. Due to the fact that this physical system of springs and electrical charges has many local minimum configurations, applying the force directed algorithm directly to a random initial layout is unlikely to give an optimal final layout. A multilevel approach can overcome this limitation. In this approach, a sequence of successively smaller graphs are generated, each captures the essential connectivity information of its parent. Global optimal layout can be found much more easily on a small graph, which are then used as a starting layout for its parent. From this initial layout, further refinement is carried out to achieve the optimal layout for the parent.

A multilevel approach has been used in many large-scale combinatorial optimization problems, such as graph partitioning [23, 30, 59], matrix ordering [36, 51], the traveling salesman problem [56], and was proved to be a very useful meta-heuristic tool [58]. A multilevel approach was later used in graph drawing [16, 25, 27, 57]. Note that a multilevel approach is not limited to the spring-electrical model, but for convenience we are introducing it in the context of this model.

A multilevel approach has three distinctive phases: coarsening, coarsest graph layout, and prolongation and refinement. In the coarsening phase, a series of coarser and coarser graphs, $G^0 = G, G^1, \dots, G^l$, are generated, each

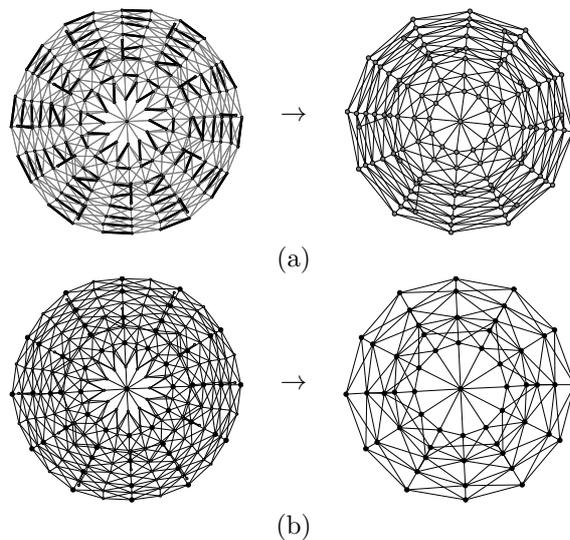


Figure 4: An illustration of graph coarsening: (a) Left: original graph with 229 vertices. Edges in a maximal independent edge set are thickened. Right: a coarser graph with 115 vertices resulted from coalescing thickened edges; (b) Left: original graph with 229 vertices. Vertices in a maximal independent vertex set are darkened. Right: a coarser graph with 55 vertices resulted from the maximal independent vertex set.

coarser graph G^{k+1} encapsulates the information needed to layout its “parent” G^k , while containing fewer vertices and edges. The coarsening continues until a graph with only a small number of vertices is reached. The optimal layout for the coarsest graph can be found cheaply. The layout on the coarser graphs are recursively prolonged to the finer graphs, with further refinement at each level. **Graph coarsening and initial layout** is the first phase in the multilevel approach. There are a number of ways to coarsen an undirected graph. One often used method is based on edge collapsing (EC) [23, 30, 59]. In this scheme, a maximal independent edge set (MIES) is selected. This is a maximal set of edges, with no edges incident to the same vertex. The vertices correspond to this edge set form a maximal matching. Each edge, and its corresponding pair of vertices, are coalesced into a new vertex. Figure 4 (a) illustrates MIES and the result of coarsening using edge collapsing.

Alternatively, coarsening can be performed based on a maximal independent vertex set (MIVS) [1]. This is a maximal set of vertices such that no two vertices in the set are connected by an edge in the graph. Edges of the coarser graph are formed through linking two vertices in the maximal independent vertex set by an edge if their distance apart is no greater than three. Figure 4 (b) illustrates MIVS and the result of coarsening using maximal independent vertex set.

Coarsest graph layout is carried out at the end of the recursive coarsening process. Coarsening is performed repeatedly until the graph is very small; at that point we can layout the graph using a suitable algorithm, for example, the force-directed Algorithm 1. Because the graph on the coarsest level is very small, it is likely that it can be laid out optimally.

The prolongation and refinement step is the third phase in a multilevel procedure. The layout on the coarser graphs are recursively interpolated to the finer graphs, with further refinement at each level.

Row “spring electrical” of Figure 5 shows drawings of two graphs using this multilevel force-directed algorithm [34]. The drawings are of good quality for both `jagmesh1`, a mesh-like graph, and `1138_bus`, a sparser graph. The spring-electrical model does suffer slightly from “warping effect” [35]. For example, for the `jagmesh1`, vertices are closer together near the boundary, compared to the interior of the mesh. This effect can be mitigated using post-processing techniques [35].

2.1.3 An open problem: more robust coarsening schemes

The multilevel approach can work efficiently, provided that the coarsening scheme is able to generate a coarsened graph with many fewer vertices than its parent graph. The aforementioned multilevel algorithm was found to work well for a lot of graphs from the graph drawing literature [34]. However, when applied to the University of Florida Sparse Matrix Collection [10], we found that, for some matrices, the coarsening scheme could not coarsen sufficiently, and the multilevel scheme has to be terminated prematurely, which results in poor drawings. An example is shown in Figure 6. On the left of the figure is the sparsity pattern of the matrix `gupta1`. From this plot we can conclude that this matrix describes a graph of three groups of vertices: those represented by the top 1/3 of the rows in the matrix plot, the middle 1/3 of the rows, and the rest. Vertices in each group are all connected to a selected few in that group; these links are seen as dense horizontal and vertical bars in the matrix plot. At the same time, vertices in the top group are connected to those in middle group, which in turn are connected to those in the bottom group, as represented by the off-diagonal lines parallel to the diagonal. However, the graph drawing in the middle of Figure 6 shows none of these structures.

This graph exemplifies many of the problematic graphs: they contain star-graph like substructures, with a lot of vertices all connected to a few vertices. Such structures pose a challenge to the usual graph coarsening schemes. These schemes are not able to coarsen graphs like that adequately. For example, Figure 7 (left) shows such a graph, with $k = 10$ vertices on the outskirts all connected to two vertices at the center. Because of this structure, any MIES can only contain two edges (the two thick edges in Figure 7). When the end vertices of the MIES are merged at the center of each edge, the resulting graph has only two fewer vertices. Therefore if k is large enough, the coarsening can be arbitrarily slow ($k/(k+2) \rightarrow 1$ as $k \rightarrow \infty$).

One solution proposed [10] is to find vertices that share the same neigh-

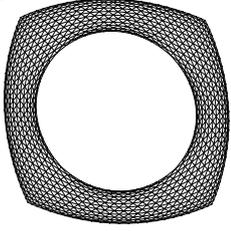
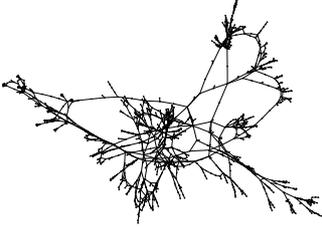
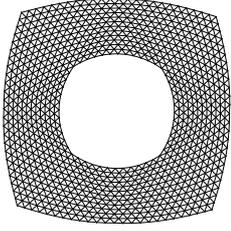
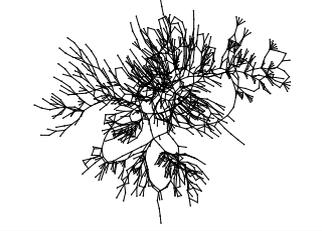
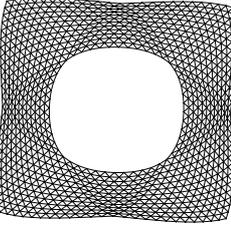
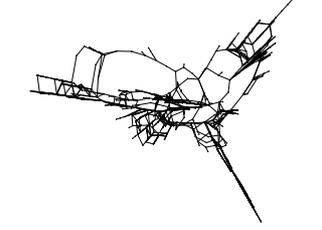
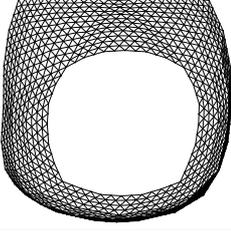
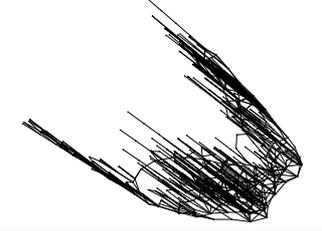
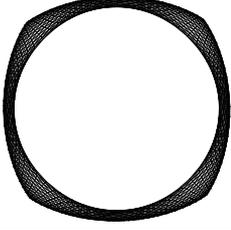
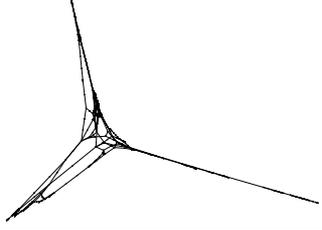
Algorithms	jagmesh1	1138_bus
spring electrical		
stress		
classical MDS		
HDE		
Hall's		

Figure 5: An overview of all algorithms described in the chapter, applied to two graphs, jagmesh1 and 1138_bus.

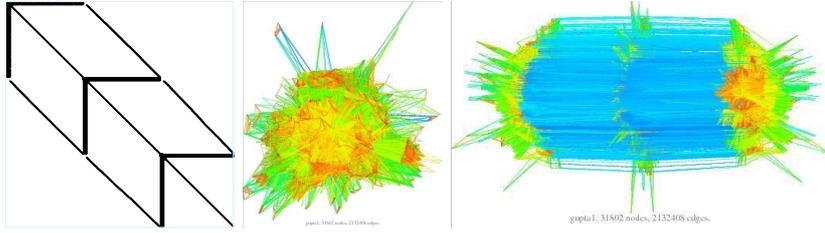


Figure 6: Matrix plot of **gupta1** matrix (left) and the initial graph drawing (middle). After applying the new coarsening scheme, the graph drawing reflects the structure of the matrix much better (right).

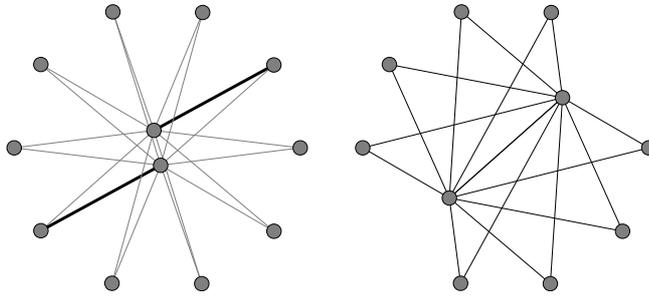


Figure 7: A maximal independent edge set based coarsening scheme fails to coarsen sufficiently a star-graph like structure: a maximal independent edge set (thick edges) (left); when merging the end vertices of the edge set at the middle of these edges, the resulting coarsened graph only has 2 fewer vertices (right).

bors. These vertices are matched in pairs. The usual MIES scheme is then used to match the remaining unmatched vertices. Finally the matched vertices are merged to get the coarsened graph. The scheme is able to overcome the slow coarsening problem associated with graphs having star-graph like substructures. Applying this scheme to the graph in Figure 8 (left) resulted in a graph with $1/2$ the number of vertices (Figure 8 right). With this new coarsening scheme, we are able to layout many more graphs aesthetically. For example, when applied to the **gupta1** matrix, the drawing at Figure 6 (right) reveals the correct visual structures as we expected, with three groups of vertices, each connected to a few within the group, and a linear connectivity relation among the groups. This new coarsening scheme is able to handle many of the graphs MIES or MIVS based coarsening schemes fail to. But a more general and robust coarsening scheme that works for all graphs is still an open problem. A promising route for investigation may be the coarsening schemes associated with algebraic multigrid (AMG) [51].

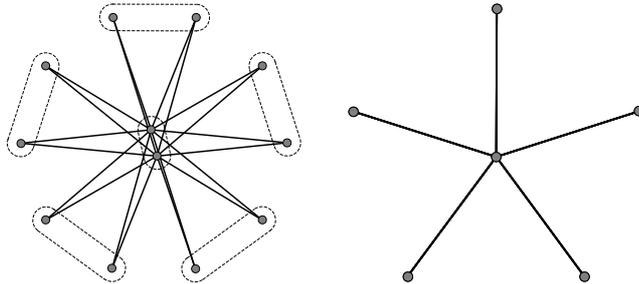


Figure 8: Matching and merging vertices with the same neighborhood structure (left, with dashed line enclosing matching vertex pairs) resulted in a new graph (right) with $1/2$ the number of vertices.

2.2 Stress and Strain Models

While the spring-electrical model is scalable and can layout graphs of millions of nodes in minutes, it does have the limitation of not coping well when edges have predefined lengths. It is possible to assign weaker attractive force and stronger repulsive force for longer edges, but such treatment is not as principled and direct as the following stress model.

2.2.1 Stress model

The stress model assumes that there are springs connecting all pairs of vertices of the graph, with the ideal spring length equal to the predefined edge length. The energy of this spring system is

$$\sum_{i \neq j} w_{ij} (\|x_i - x_j\| - d_{ij})^2, \quad (2)$$

where d_{ij} is the ideal distance between vertices i and j , and w_{ij} is a weight factor, typically $1/d_{ij}^2$. The layout that minimizes the above stress energy is an optimal layout of the graph according to this model.

The stress model has its root in Multidimensional Scaling (MDS) [40, 41]. Note that typically we only know the ideal distance between vertices that share an edge, which is usually taken to be one for graphs without predefined edge length. Alternatively, it has been proposed to set the edge length equal the total number of non-common neighbors of the two end vertices [19]. For other vertex pairs, one way to define d_{ij} is to take it as the shortest distance between vertex i and j . The practice of taking the shortest graph distance as the ideal edge length date back at least to 1980 in social network layout [6], and in graph drawing using classical MDS[41], but is often attributed to Kamada and Kawai [37].

There are several ways to minimize (2). A force-directed algorithm (Algorithm 1) can be used. The repulsive/attractive force exerted on vertex i from

the spring between vertices i and j is

$$F(i, j) = -w_{ij}(\|x_i - x_j\| - d_{ij}) \frac{x_i - x_j}{\|x_i - x_j\|}, \quad i \neq j. \quad (3)$$

Stress-majorization: a stress-majorization technique [19] can be employed to solve the stress model. Consider the cost function (2),

$$\sum_{i \neq j} w_{ij} (\|x_i - x_j\| - d_{ij})^2 = \sum_{i \neq j} (w_{ij} \|x_i - x_j\|^2 - 2d_{ij} w_{ij} \|x_i - x_j\| + w_{ij} d_{ij}^2)$$

On the right hand side, the first and third terms are either constant or quadratic with regard to x , except the second one. Using the Cauchy-Schwartz inequality, $(x_i - x_j)^T (y_i - y_j) \leq \|x_i - x_j\| \|y_i - y_j\|$, we can bound the cost function by

$$g(x, y) = \sum_{i \neq j} \left(w_{ij} \|x_i - x_j\|^2 - 2d_{ij} w_{ij} \frac{(x_i - x_j)^T (y_i - y_j)}{\|y_i - y_j\|} + w_{ij} d_{ij}^2 \right),$$

with the bound tight when $y = x$. The idea of stress majorization is to minimize a sequence of quadratic function $g(x, y^k)$, with $y^0 = x^0$ the initial layout, and subsequent y^k the result of minimizing $g(x, y^{k-1})$, $k = 1, 2, \dots$

The minimum of the quadratic function $g(x, y)$ is derived by setting $\partial_{x_i} g(x, y) = 0$, giving

$$L_w x = L_{w,d} y \quad (4)$$

where the weighted Laplacian matrix L_w has elements

$$(L_w)_{ij} = \begin{cases} \sum_{i \neq l} w_{il}, & i = j \\ -w_{ij}, & i \neq j \end{cases}$$

and matrix $L_{w,d}$ has elements

$$(L_{w,d})_{ij} = \begin{cases} \sum_{i \neq l} w_{il} d_{il} / \|y_i - y_l\|, & i = j \\ -w_{ij} d_{ij} / \|y_i - y_j\|, & i \neq j \end{cases}$$

In summary, the process of finding a minima of (2) becomes that of solving a series of linear systems (4), with the solution x served as y in the next iteration. This iterative process is found to be quite robust, although for large graphs, it still benefits from a good initial layout. Row “stress” of Figure 5 gives drawings of the stress model. It performed very well on both graphs.

2.2.2 Strain model (classical MDS)

The strain model, also known as classical MDS [53], predates the stress model. Classical MDS tries to fit the inner product of positions, instead of the distance between points. Specifically, assume that the final embedding is centered around the origin:

$$\sum_{i=1}^{|V|} x_i = 0.$$

Furthermore, assume that in the ideal case, the embedding fits the distance exactly:

$$\|x_i - x_j\| = d_{ij}. \quad (5)$$

It is then easy to prove [61] that the product of the positions, $b_{ij} = x_i^T x_j$, can be expressed as the squared and double centered distance,

$$b_{ij} = x_i^T x_j = -1/2 \left(d_{ij}^2 - \frac{1}{|V|} \sum_{k=1}^{|V|} d_{kj}^2 - \frac{1}{|V|} \sum_{k=1}^{|V|} d_{ik}^2 + \frac{1}{|V|^2} \sum_{k=1}^{|V|} \sum_{l=1}^{|V|} d_{kl}^2 \right). \quad (6)$$

In real data, it is unlikely that we can find an embedding that fits the distances perfectly, hence assumption (5) does not stand. But we would still expect that b_{ij} is a good approximation of $x_i^T x_j$. Therefore we try to find an embedding that minimizes the difference between the two,

$$\min_X \|X^T X - B\|_F, \quad (7)$$

where X is the $|V| \times d$ dimensional matrix of x_i 's, B is the $|V| \times |V|$ symmetric matrix of b_{ij} 's, and $\|\cdot\|_F$ is the Frobenius norm. If the eigen-decomposition of B is $B = Q^T \Lambda Q$, then the solution to (7) is $X = \Lambda_d^{1/2} Q$, where Λ_d is the diagonal matrix of Λ , with all but the d largest eigenvalues on the diagonal set to zero. Because the strain model does not fit the distance directly, graph drawings given by solving this model are not as satisfactory as those using the stress model [6], but it can be used as a good starting point for the stress model. Row “classical MDS” of Figure 5 gives drawings using classical MDS. It performed well on the mesh like `jagmesh1` graph, but not so well on the sparser `1138_bus` graph, where vertices cling close to each other, making some details of the graph unclear.

2.2.3 MDS for large graphs

In a typical usage of the stress or strain models, the ideal distance between all pairs of vertices has to be calculated, which requires an all-pair shortest path calculation. Using Johnson’s algorithm, this needs $O(|V|^2 \log |V| + |V||E|)$ computation, and a storage of $O(|V|^2)$. Therefore for very large graphs, this formulation is computationally expensive and memory prohibitive. A number

of strategies [5, 11, 19] have been proposed to approximately minimize (2) or (7).

A **multiscale algorithm** [16, 28] applies the multilevel approach in solving the stress model. In the GRIP algorithm [16], graph coarsening is carried out through vertex filtration, an idea similar to the maximal independent vertex set. A sequence of vertex sets, $V^0 = V \subset V^1 \subset V^2, \dots, \subset V_L$, is generated. However, coarser graphs are not constructed explicitly. Instead, a vertex set V^k at level k of the vertex set hierarchy is constructed so that distance between vertices is at least $2^{k-1} + 1$. On each level k , the stress model is solved by a force-directed procedure. The spring force (3) on each vertex $i \in V^k$ is calculated by considering a neighborhood $N^k(i)$ of this vertex, with $N^k(i)$ the set of vertices in level k , chosen so that the total number of vertices in this set is $O(|E|/|V^k|)$. Thus the force calculation on each level can be done in time $O(|E|)$. It was proved that with this multilevel procedure and the localized force calculation algorithm, for a graph of bounded degree, the algorithm has close to linear computational and memory complexity.

LandmarkMDS [11] approximates the result of classical MDS by choosing $k \ll |V|$ vertices as landmarks, and calculating a layout of these vertices using the classical MDS, based on distances among these vertices. The k landmarks are chosen to be well dispersed in the graph. One possibility is to use a MaxMin strategy where the first landmark is randomly chosen, then each subsequent landmark is chosen to be furthest away from the previous landmarks, which is a well-known 2-approximation to the k -center problem. The positions of the rest of vertices are then calculated by placing these vertices at the weighted barycenter, with weights based on distances to the landmarks. So essentially classical MDS is applied to a $k \times k$ submatrix of the $|V| \times |V|$ matrix B . The complexity of this algorithm is $O(k|V| + k^2)$, and only $O(k|V|)$ distances need to be stored.

PivotMDS [5], on the other hand, takes a $|V| \times k$ submatrix C of B . The two eigenvectors corresponding to the largest eigenvalues of the $|V| \times |V|$ matrix CC^T are then calculated using power iterations and used as the x - and y - coordinates. By an algebraic argument, if v is an eigenvector of the $k \times k$ matrix $C^T C$, then

$$CC^T(Cv) = C(C^T Cv) = \lambda(Cv),$$

hence Cv is an eigenvector of CC^T . Therefore PivotMDS proceeds by finding the largest eigenvectors of the smaller $k \times k$ matrix $C^T C$, then projects back to $|V|$ -dimensional space by multiplying them with C . Using this technique, the overall complexity is similar to LandmarkMDS, but unlike LandmarkMDS, PivotMDS utilizes distances between landmark vertices and other vertices in the matrix product. In practice PivotMDS was found to give drawings that are closer to the classical MDS than LandmarkMDS, and both are very efficient when used with a small number (e.g., $k \approx 50$) of pivots/landmarks.

It is worth pointing out that, for sparse graphs, there is a limitation in both algorithms. For example, if the graph is a tree, and pivots/landmarks are

chosen to be non-leaves, then two leaf nodes that have the same parent will have exact the same distances to any of the pivots/landmarks, consequently their final positions based on these algorithms will also be the same. This problem may be alleviated by utilizing the layout given by these algorithms as an initial placement for a stress model based algorithm, but taking only a sparse set of terms in the stress function [6, 19].

2.3 High-Dimensional Embedding

The high-dimensional embedding (HDE) algorithm [29] finds coordinates of vertices in a k -dimensional space, then projects back to two or three dimensional space.

First, a k -dimensional coordinate system is created based on k -centers, where k -centers are chosen as in LandmarkMDS/PivotMDS (Section 2.2.3). The graph distances from each vertex to the k -centers form a k -dimensional coordinate system. The $|V|$ coordinate vectors form an $|V| \times k$ matrix Y , where the i -th row of Y is the k -dimensional coordinates for vertex i .

Since it is only possible to visualize in two or three dimensions, and since the coordinates may be correlated, the coordinates are projected back to two or three dimensions by suitable linear combinations that minimize correlations. To make this projection shift-invariant, Y is first normalized so that the center of gravity of the vertices is at the origin, i.e.,

$$Y := Y - \frac{1}{|V|}ee^TY,$$

where e is the $|V|$ -dimensional vector of all 1's.

Assume we want to project the k dimensional coordinate system into 2-dimensions. Let v_1 and v_2 be two k -dimensional linear combination vectors, so that Yv_1 and Yv_2 form the x - and y - coordinates. The two linear combinations should be uncorrelated, so we take them to be orthogonal to each other:

$$v_1^TY^TYv_2 = 0.$$

Furthermore, each should be as far away from 0 as possible, so we want to maximize

$$\frac{v_l^TY^TYv_l}{\|v_l\|}, \quad l = 0, 1.$$

These can be achieved by taking v_1 and v_2 to be the two eigenvectors that correspond to the two largest eigenvalues of the $k \times k$ symmetric matrix Y^TY . This process of choosing highly uncorrelated vectors out of high dimensional data is known as principal component analysis. The final x - and y - coordinates are given by Yv_1 and Yv_2 .

HDE clearly has many commonalities to PivotMDS. Each utilizes k -centers, and each finds the largest eigenvectors of a $k \times k$ dimensional matrix derived by multiplying the transpose of a $|V| \times k$ matrix with itself. The main difference

is that in high-dimensional embedding this $|V| \times k$ matrix consists of distances to the k -centers, while in PivotMDS this matrix consists of distances squared and double centered (see (6)). High-dimensional embedding suffers from the same limitation as PivotMDS and LandmarkMDS, in that it is not able to layout sparse graphs as well as mesh-like graphs. Row “HDE” of Figure 5 gives drawings using HDE. It performs particularly badly on `1138_bus` graph, with vertices close to each other, obscuring many details.

2.4 Hall’s Algorithm

In 1970, Hall [26] remarked that many sequencing and placement problems could be characterized as finding locations of points which minimize the weighted sum of square distances. In our notation, what he proposed was to minimize

$$\sum_{i \leftrightarrow j} w_{ij} \|x_i - x_j\|^2, \text{ subject to } \sum_{k=1}^{|V|} x_k^2 = 1$$

where x_i is the 1-dimensional coordinate value for vertex i . The objective function can be written as

$$\sum_{i \leftrightarrow j} w_{ij} \|x_i - x_j\|^2 = x^T L_w x,$$

with $x = \{x_1, x_2, \dots, x_{|V|}\}$. Here L_w is the weighted Laplacian matrix with elements

$$(L_w)_{ij} = \begin{cases} \sum_{(i,l) \in E} w_{il}, & i = j \\ -w_{ij}, & i \neq j \end{cases}$$

For a connected graph, the Laplacian is positive semi-definite with one eigenvalue of 0 corresponding to the trivial eigenvector of all 1’s. The solution x of the minimization problem is the eigenvector corresponding to the smallest positive eigenvalue of the weighted Laplacian L_w . We can achieve a 2-dimensional layout by taking the two eigenvectors corresponding to the two smallest positive eigenvalues. Row “Hall’s” of Figure 5 gives drawings employing Hall’s algorithm. It performed reasonably well on the mesh like `jagmesh1` graph, but is almost useless on the sparser `1138_bus` graph.

Koren et al. [38] proposed an extremely fast algorithm for calculating the two extreme eigenvectors using a multilevel algorithm. The algorithm is called ACE (Algebraic multigrid Computation of Eigenvectors). Using this algorithm, they were able to layout graphs of millions of nodes in less than a minute. However, the fundamental weakness of Hall’s algorithm on sparse graphs remains.

3 Examples of Large Graph Drawings

In this section we give example drawings of some large graphs. These graphs are drawn using a multilevel spring-electrical model based code, `sfdp` [34], available

from `graphviz` [21].

One rich source of large graphs is the University of Florida Sparse Matrix Collection [10]. This is a collection of 2272 (as of January, 2010) sparse matrices. The largest matrix has 27 million rows and columns. Graph visualization provides a way to visualize these matrices and get a glimpse of the application underneath these matrices. Figure 9 gives the drawing of four graphs in this collection. For instance, by its name, the `boneS10` graph could be from a model of the bone. The porous structure can be seen clearly from the drawing. More drawings for all the matrices in the collection can be found at [33].

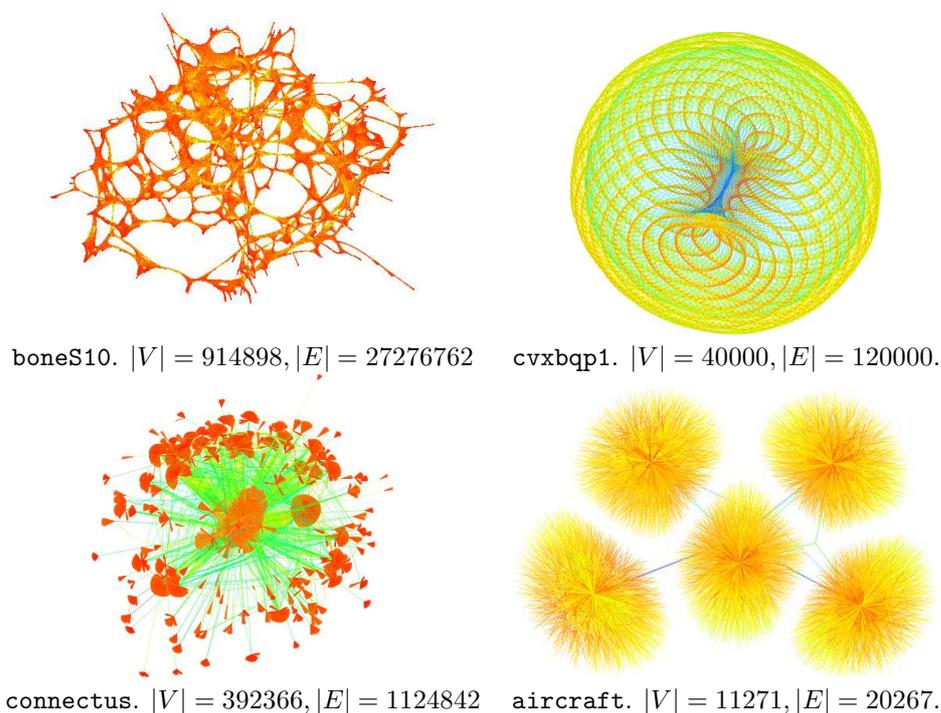


Figure 9: Drawing of some graphs from the University of Florida Sparse Matrix Collection [10]. More drawings can be found at [33].

Figure 10 gives a drawing of the tree of life. The data came from the Tree of Life Project [43]. This project documents phylogeny of organisms, i.e., the history of organismal lineages as they change through time. It implies that different species arise from previous forms via descent, and that all organisms are connected by the passage of genes along the branches of the phylogenetic tree. The data, as collected in February, 2009, contains 76425 species, representing a tiny fraction of the estimated 5 to 100 million species on Earth today. Figure 10 clearly shows the tree structure. The root vertex “Life on Earth” (see

Acknowledgments

The author would like to thank Emden Gansner and Stephen North, as well as anonymous referees, for valuable comments.

References

- [1] S. T. Barnard and H. D. Simon. Fast multilevel implementation of recursive spectral bisection for partitioning unstructured problems. *Concurrency: Practice and Experience*, 6:101–117, 1994.
- [2] J. Barnes and P. Hut. A hierarchical $O(N \log N)$ force-calculation algorithm. *Nature*, 324:446–449, 1986.
- [3] V. Batagelj. Visualization of large networks. In R. A. Meyers, editor, *Encyclopedia of Complexity and Systems Science*. Springer, New York, 2009.
- [4] G. D. Battista, P. Eades, R. Tamassia, and I. G. Tollis. *Algorithms for the visualization of Graphs*. Prentice-Hall, 1999.
- [5] U. Brandes and C. Pich. Eigensolver methods for progressive multidimensional scaling of large data. In *Proc. 14th Intl. Symp. Graph Drawing (GD '06)*, volume 4372 of *LNCS*, pages 42–53, 2007.
- [6] U. Brandes and C. Pich. An experimental study on distance based graph drawing. In *Proc. 16th Intl. Symp. Graph Drawing (GD '08)*, volume 5417 of *LNCS*, pages 218–229. Springer-Verlag, 2009.
- [7] I. Brass and A. Frick. Fast interactive 3-D graph visualization. *LNCS*, 1027:99–11, 1995.
- [8] A. Burton, A. J. Field, and H. W. To. A cell-cell Barnes Hut algorithm for fast particle simulation. *Australian Computer Science Communications*, 20:267–278, 1998.
- [9] C. Collins, G. Penn, and S. Carpendale. Bubble sets: Revealing set relations with isocontours over existing visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1009–1016, 2009.
- [10] T. A. Davis and Y. F. Hu. University of florida sparse matrix collection. *ACM Transaction on Mathematical Software*, 2011 (to appear).
- [11] V. de Solva and J. B. Tenenbaum. Global versus local methods in nonlinear dimensionality reduction. In *Advances in Neural Information Processing Systems 15*, pages 721–728. MIT Press, 2003.
- [12] P. Eades. A heuristic for graph drawing. *Congressus Numerantium*, 42:149–160, 1984.
- [13] L. Euler. Commentarii academiae scientiarum petropolitanae. *Solutio problematis ad geometriam situs pertinentis*, 8:128–140, 1741.
- [14] Y. Frishman and A. Tal. Online dynamic graph drawing. In *proceeding of Eurographics/IEEE VGTC Symposium on Visualization (EuroVis)*, pages 75–82, 2007.

- [15] T. M. J. Fruchterman and E. M. Reingold. Graph drawing by force directed placement. *Software - Practice and Experience*, 21:1129–1164, 1991.
- [16] P. Gajer, M. T. Goodrich, and S. G. Kobourov. A fast multi-dimensional algorithm for drawing large graphs. *LNCS*, 1984:211 – 221, 2000.
- [17] E. R. Gansner, Y. F. Hu, and S. G. Kobourov. Gmap: Visualizing graphs and clusters as map. In *Proceedings of IEEE Pacific Visualization Symposium*, pages 201 – 208, 2010.
- [18] E. R. Gansner, Y. Koren, and S. North. Topological fisheye views for visualizing large graphs. *IEEE Transactions on Visualization and Computer Graphics*, 11:457–468, 2005.
- [19] E. R. Gansner, Y. Koren, and S. C. North. Graph drawing by stress majorization. In *Proc. 12th Intl. Symp. Graph Drawing (GD '04)*, volume 3383 of *LNCS*, pages 239–250. Springer, 2004.
- [20] E. R. Gansner, E. Koutsofios, S. C. North, and K. P. Vo. A technique for drawing directed graphs. *IEEE Trans. Softw. Eng.*, 19:214–230, 1993.
- [21] E. R. Gansner and S. North. An open graph visualization system and its applications to software engineering. *Software - Practice & Experience*, 30:1203–1233, 2000.
- [22] L. F. Greengard. *The rapid evaluation of potential fields in particle systems*. The MIT Press, Cambridge, Massachusetts, 1988.
- [23] A. Gupta, G. Karypis, and V. Kumar. Highly scalable parallel algorithms for sparse matrix factorization. *IEEE Transactions on Parallel and Distributed Systems*, 5:502–520, 1997.
- [24] S. Hachul and M. Jünger. Drawing large graphs with a potential field based multilevel algorithm. In *Proc. 12th Intl. Symp. Graph Drawing (GD '04)*, volume 3383 of *LNCS*, pages 285–295. Springer, 2004.
- [25] R. Hadany and D. Harel. A multi-scale algorithm for drawing graphs nicely. *Discrete Applied Mathematics*, 113:3–21, 2001.
- [26] K. M. Hall. An r -dimensional quadratic placement algorithm. *Management Science*, 17:219–229, 1970.
- [27] D. Harel and Y. Koren. A fast multi-scale method for drawing large graphs. *J. Graph Algorithms and Applications*, 6:179–202, 2002.
- [28] D. Harel and Y. Koren. A fast multi-scale method for drawing large graphs. *Journal of graph algorithms and applications*, 6:179–202, 2002.
- [29] D. Harel and Y. Koren. Graph drawing by high-dimensional embedding. *lncs*, pages 207–219, 2002.

- [30] B. Hendrickson and R. Leland. A multilevel algorithm for partitioning graphs. Technical Report SAND93-1301, Sandia National Laboratories, Albuquerque, NM, 1993. Also in Proceeding of Supercomputing'95 (http://www.supercomp.org/sc95/proceedings/509_BHEN/SC95.HTM).
- [31] I. Herman, G. Melancon, and M. S. Marshall. Graph visualization and navigation in information visualization: A survey. *IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS*, 6:24–43, 2000.
- [32] D. Holten and J. J. van Wijk. Force-directed edge bundling for graph visualization. *Computer Graphics Forum*, 28:983–990, 2009.
- [33] Y. F. Hu. A gallery of large graphs. <http://research.att.com/~yifanhu/GALLERY/GRAPHS/index.html>.
- [34] Y. F. Hu. Efficient and high quality force-directed graph drawing. *Mathematica Journal*, 10:37–71, 2005.
- [35] Y. F. Hu and Y. Koren. Extending the spring-electrical model to overcome warping effects. In *Proceedings of IEEE Pacific Visualization Symposium*, pages 129–136. IEEE Computer Society, 2009.
- [36] Y. F. Hu and J. A. Scott. A multilevel algorithm for wavefront reduction. *SIAM Journal on Scientific Computing*, 23:1352–1375, 2001.
- [37] T. Kamada and S. Kawai. An algorithm for drawing general undirected graphs. *Information Processing Letters*, 31:7–15, 1989.
- [38] Y. Koren, L. Carmel, and D. Harel. Ace: A fast multiscale eigenvectors computation for drawing huge graphs. In *INFOVIS '02: Proceedings of the IEEE Symposium on Information Visualization (InfoVis'02)*, pages 137–144, Washington, DC, USA, 2002. IEEE Computer Society.
- [39] E. Krujaa, J. Marks, A. Blair, and R. Waters. A short note on the history of graph drawing. In *Proc. 9th Intl. Symp. Graph Drawing (GD '01)*, pages 272–286. Springer-Verlag, London, UK, 2002.
- [40] J. B. Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29:1–27, 1964.
- [41] J. B. Kruskal and J. B. Seery. Designing network diagrams. In *Proceedings of the First General Conference on Social Graphics*, pages 22–50, Washington, D.C., July 1980. U. S. Department of the Census. Bell Laboratories Technical Report No. 49.
- [42] J. Lamping, R. Rao, and P. Pirolli. A focus+context technique based on hyperbolic geometry for visualizing large hierarchies. In *SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS (CHI '95)*, pages 401–408. ACM, 1995.

- [43] D. R. Maddison and K.-S. S. (eds.). The tree of life web project. <http://tolweb.org>, 2007.
- [44] T. Moscovich, F. Chevalier, N. Henry, E. Pietriga, and J. Fekete. Topology-aware navigation in large networks. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*, pages 2319–2328, New York, NY, USA, 2009. ACM.
- [45] T. Munzner. Exploring large graphs in 3d hyperbolic space. *IEEE Comput. Graph. Appl.*, 18:18–23, 1998.
- [46] T. Munzner and P. Burchard. Visualizing the structure of the world wide web in 3d hyperbolic space. In *VRML '95: Proceedings of the first symposium on Virtual reality modeling language*, pages 33–38, New York, NY, USA, 1995. ACM.
- [47] A. Noack. An energy model for visual graph clustering. In *Proceedings of the 11th International Symposium on Graph Drawing (GD 2003)*, volume 2912 of *LNCS*, pages 425–436. Springer, 2004.
- [48] S. Pfalzner and P. Gibbon. *Many-Body Tree Methods in Physics*. Cambridge University Press, Cambridge, 1996.
- [49] A. Quigley. *Large scale relational information visualization, clustering, and abstraction*. PhD thesis, Department of Computer Science and Software Engineering, University of Newcastle, Australia, 2001.
- [50] A. Quigley and P. Eades. Fade: Graph drawing, clustering, and visual abstraction. *LNCS*, 1984:183–196, 2000.
- [51] I. Safro, D. Ron, and A. Brandt. Multilevel algorithms for linear ordering problems. *J. Exp. Algorithmics*, 13:1.4–1.20, 2009.
- [52] K. Sugiyama, S. Tagawa, and M. Toda. Methods for visual understanding of hierarchical systems. *IEEE Trans. Systems, Man and Cybernetics*, SMC-11(2):109–125, 1981.
- [53] W. S. Torgerson. Multidimensional scaling: I. theory and method. *Psychometrika*, 17:401–419, 1952.
- [54] D. Tunkelang. *A Numerical Optimization Approach to General Graph Drawing*. PhD thesis, Carnegie Mellon University, 1999.
- [55] W. Tutte. How to draw a graph. *Proceedings of the London Mathematical Society*, 13:743–768, 1963.
- [56] C. Walshaw. A multilevel approach to the travelling salesman problem. *Oper. Res.*, 50:862–877, 2002.
- [57] C. Walshaw. A multilevel algorithm for force-directed graph drawing. *J. Graph Algorithms and Applications*, 7:253–285, 2003.

- [58] C. Walshaw. Multilevel refinement for combinatorial optimisation problems. *Annals of Operations Research*, pages 325–372, 2004.
- [59] C. Walshaw, M. Cross, and M. G. Everett. Parallel dynamic graph partitioning for adaptive unstructured meshes. *Journal of Parallel and Distributed Computing*, 47:102–108, 1997.
- [60] D. Watts and S. Strogatz. Collective dynamics of “small-world” networks. *Nature*, 393:440–442, 1998.
- [61] G. Young and A. S. Householder. Discussion of a set of points in terms of their mutual distances. *Psychometrika*, 3:19–22, 1938.