More Flexible Radial Layout

Ulrik Brandes and Christian Pich

1 Department of Computer & Information Science, University of Konstanz
   Ulrik.Brandes@uni-konstanz.de
2 Chair of Systems Design, ETH Zürich
   cpich@ethz.ch

Abstract. We describe an algorithm for radial layout of undirected graphs, in which nodes are constrained to the circumferences of a set of concentric circles around the origin. Such constraints frequently occur in the layout of social or policy networks, when structural centrality is mapped to geometric centrality, or when the primary intention of the layout is the display of the vicinity of a distinguished node. We extend stress majorization by a weighting scheme which imposes radial constraints on the layout but also tries to preserve as much information about the graph structure as possible.

1 Introduction

In radial graph layout the nodes are constrained to be located on a set of concentric rings; for some or all nodes in a graph a radius is given, which typically encodes the results of a preceding analysis step. Such drawings date back to the 1940s and are called ring or target diagrams [18]. The interpretation of these rings is specific to the particular application at hand. The overall goals which guide the design of radial layouts can be expressed as two criteria and are possibly contradictory:

– Representation of distances: The Euclidean distance between two nodes in the drawing should correspond to their graph-theoretical distance. This is a general objective common to all “organic” layout styles.
– Radial constraints: Nodes are associated with the radius of a circle centered in the origin, and are constrained to be placed on the circumference of this circle.

Radial layout occurs as a task in several applications. It is used for the exploration of large hierarchies in [19]; the hierarchy is laid out radially as a tree, followed by an incremental force-based placement. This approach was later modified for dynamic real-time exploration of a filesharing network in [20], where users interactively select a node to be moved into the center, while the current immediate surrounding of that node is updated.

† Part of this work was done while the author was at the University of Konstanz, Department of Computer & Information Science.
A different approach is to extend the definition of level planarity to discrete radial levels [1]; the traditional Sugiyama framework is enhanced accordingly for linear-time embedding of level-planar graphs.

In the case of continuous radii representing some kind of substance, unary constraints are imposed on the drawing for mapping centrality scores to visual centrality [4]. The layout is done by simulated annealing, which allows for penalty costs, e.g. for edge crossings, and is very flexible but also computationally demanding; this prohibits interactivity even for moderately sized graphs.

In the following, the two essential goals above are explicitly formulated as objective functions, which measure how far a layout is from meeting the criteria, and which are sought to be minimized. While the first objective is captured by the traditional energy or stress measures, we try to fulfill the second objective by introducing radial constraints into the energy-based layout model and using a linear combination of the two objectives.

Quite recently, extensions of the stress term have been used for drawing graphs with explicitly formulated aesthetic criteria, such as the uniform scattering of the nodes in a graph over a unit disk [16], penalizing node overlaps [11], or preserving a given topology [10].

All these approaches modify the distances themselves in one form or another, while the approach presented in this contribution is based on engineering the weights used in the stress minimization model. The weights are coefficients of error terms involved in the quality criteria to be minimized. If chosen carefully, the weights can be used to influence the configuration resulting from optimizing the stress function modified by these weights; see Fig. 1 for an example. We are not aware of previous work which makes systematic use of such a weighting scheme to take up a particular perspective on a data set.

![Layouts of a social network (2075 nodes, 4769 edges), consisting of two known clusters. The brightness of node colors is proportional to the graph-theoretical distances from a distinguished focal node, which also defines the radii used in the constrained layout, in which the two clusters are still visible.](image)

**Fig. 1.** Layouts of a social network (2075 nodes, 4769 edges), consisting of two known clusters. The brightness of node colors is proportional to the graph-theoretical distances from a distinguished focal node, which also defines the radii used in the constrained layout, in which the two clusters are still visible.
2 Preliminaries

Let $G = (V, E)$ be an undirected graph, $E \subseteq \binom{V}{2}$. We will denote the cardinalities of the node and edge sets by $n = |V|$ and $m = |E|$, respectively; it will be convenient to index nodes by numbers, $V = \{1, \ldots, n\}$. The graph-theoretical distance between two nodes $i, j$ is the number of edges on a shortest path between $i$ and $j$ and is denoted $d_{ij}$ or, when there is no danger of confusion, $d_{ij}$. The matrix $D = (d_{ij})_{ij} \in \mathbb{R}^{n \times n}$ contains the distances between every two nodes in $G$; the diameter of $G$ is the maximum distance between any two nodes in $G$, $\text{diam}(G) = \max_{i,j \in V} d_{ij}$. All graphs are assumed to be connected; otherwise, all connected components are just considered individually.

Two-dimensional node positions are denoted $p(i) = (x_i, y_i)$; a layout of $n$ nodes is captured by two column vectors $x = [x_1, \ldots, x_n]^T, y = [y_1, \ldots, y_n]^T \in \mathbb{R}^n$. The Euclidean distance between two nodes in a given layout $p$ is defined as $\|p(i) - p(j)\| = (|x_i - x_j|^2 + |y_i - y_j|^2)^{1/2}$.

3 Stress, Weights, and Constraints

3.1 Stress

The foundation of the method presented in the following is multidimensional scaling (MDS) [2,7]; originating in psychometrics and the social sciences, MDS has been established and widely used in the graph drawing community for more than three decades, as energy-based placement [15]. While there is a wide range of variants and extensions, we will concentrate on the stress minimization approach [12] in this contribution.

Given a set of desired distances among a set of $n$ objects, the overall goal is to place these objects in a low-dimensional Euclidean space in such a way that the resulting distances fit the desired ones as well as possible. In the graph drawing literature, the desired distances are usually graph-theoretical distances $d_{ij}$, and the goal is to find two-dimensional positions $p(1), \ldots, p(n)$ with

$\|p(i) - p(j)\| \approx d_{ij}$  \hspace{1cm} (1)

attained as closely as possible for all pairs $i, j$. When the configuration is not required to satisfy any further constraints, the objective function, called stress, is the sum of squared residuals

$\sigma(p) = \sum_{i<j} w_{ij} (d_{ij} - \|p(i) - p(j)\|)^2$  \hspace{1cm} (2)

over all the $n(n-1)/2$ pairs of nodes, where $w_{ij} \geq 0$ is a weight for the contribution of the particular error term $(d_{ij} - \|p(i) - p(j)\|)^2$ for pair $i, j$ to the stress.

There is a wide consensus that configurations with a small stress value tend to be aesthetically pleasing. The state-of-the-art approach to finding such layouts
is stress majorization \[8,12\]; starting from an initial configuration, it generates a sequence of improving layouts. When no geometric coordinates are at hand, the iteration may be initialized at random; however, more favorable and robust strategies are available for initial layouts, such as classical MDS \[5\].

In an iterative process, new coordinates \( \hat{x} = [\hat{x}_1, \ldots, \hat{x}_n]^T, \hat{y} = [\hat{y}_1, \ldots, \hat{y}_n]^T \in \mathbb{R}^n \) are computed from the current ones with the update

\[
\hat{x}_i \leftarrow \frac{\sum_{j \neq i} w_{ij} \left( x_j + d_{ij} \cdot (x_i - x_j) \cdot b_{ij} \right)}{\sum_{j \neq i} w_{ij}} \quad (3)
\]

\[
\hat{y}_i \leftarrow \frac{\sum_{j \neq i} w_{ij} \left( y_j + d_{ij} \cdot (y_i - y_j) \cdot b_{ij} \right)}{\sum_{j \neq i} w_{ij}} \quad (4)
\]

where

\[
b_{ij} = \begin{cases} 
\frac{1}{\|p(i) - p(j)\|} & \text{if } \|p(i) - p(j)\| > 0, \\
0 & \text{otherwise.} 
\end{cases} \quad (5)
\]

This is repeated until the relative change in the configuration is below some threshold value, or after a predefined number of steps. The sequence of layouts generated in this way can be shown to have non-increasing stress and to converge towards a local minimum \[9\].

### 3.2 Weights for Constraints

In early applications of MDS, each pair \( i, j \) of objects was assigned the same unit weight by setting \( w_{ij} = 1 \) in (2); when a desired distance was not known for a pair, this pair was simply ignored by using a zero weight for its contribution to the stress.

In graph drawing it is a de-facto standard to set \( w_{ij} = d_{ij}^{-2} \) to emphasize the quality of the fit of local distances, i.e., the contribution of pairs \( i, j \) with smaller target distances is increased compared to pairs with larger distances. This weighting scheme was introduced in elastic scaling by McGee \[17\], and is equal to the one used by Kamada and Kawai \[15\]. Instead of fitting absolute values by minimizing absolute residual error terms \((d_{ij} - \|p(i) - p(j)\|)^2\), the goal is to achieve a fit of the distance magnitudes, expressed by relative error terms \((1 - \|p(i) - p(j)\|/d_{ij})^2\). Summing these over all pairs gives the sum

\[
\sum_{i < j} \left(1 - \frac{\|p(i) - p(j)\|}{d_{ij}}\right)^2 = \sum_{i < j} \frac{1}{d_{ij}^2} (d_{ij} - \|p(i) - p(j)\|)^2. \quad (6)
\]

In this sum the impact of larger distances in the unweighted stress (2) is lessened, which is due to the square in the error term.

A reason for the favorable aesthetic properties of low-stress layouts is that no node is preferred over others because minimizing the objective function tries to achieve a balance in the fit of the desired distances. In most scenarios this is appropriate and tends to give the drawing a balanced appearance.
In some cases, users want to put more emphasis on some nodes, while other nodes are regarded less important, by centering the view on a node and visualizing this node’s neighborhood more prominently. This can be done by introducing suitable constraints on the configuration; when these constraints can be formulated as desired distances, choosing the weights in a suitable way allows for imposing them on the resulting layout.

What follows is a general framework for constraining drawings; while the range of possible applications is wide, our contribution will concentrate on the radial scenario. To avoid confusion, the objective function (2) will be termed distance stress and denoted $\sigma_W(p)$, with the subscript indicating that the stress is modified by a weight matrix $W = (w_{ij})_{ij} \in \mathbb{R}^{n \times n}$. This stress model is enhanced by a second set of weights $Z = (z_{ij})_{ij}$ used in the constraint stress

$$\sigma_Z(p) = \sum_{i<j} z_{ij} (d_{ij} - \|p(i) - p(j)\|)^2$$  

whose minimization tries to fit the same distances and hence aims at representing the same information, but highlights different aspects.

### 3.3 Interpolated Weights

A straightforward approach to imposing the constraints expressed in a weight matrix is to directly minimize (7), but the resulting solutions tend to be trivial; for example, consider a linear layout in which $x_i = r_i, y_i = 0$ for all $i \in V$. Instead, it is more effective to combine distance and constraint stress into a joint majorization process, operating on a linear combination of the stress measures $\sigma_W(p)$ and $\sigma_Z(p)$.

Initially, the nodes are allowed to move freely without considering the constraints at all, by minimizing just $\sigma_W(x, y)$. Then, the constraints are granted more and more control over the appearance of the drawing by dynamically changing the coefficients in this combination, and the bias is shifted from one to the other criterion [3]. The influences of the distance and the radial components are determined by the coefficients in the linear combination

$$\sigma_t(p) = (1 - t) \cdot \sigma_W(p) + t \cdot \sigma_Z(p)$$

and the update terms for the majorization process (3) and (4) become

$$\dot{x}_i \leftarrow \frac{\sum_{j \neq i} ((1 - t) \cdot w_{ij} + t \cdot z_{ij}) \cdot (x_j + d_{ij} \cdot (x_i - x_j) \cdot b_{ij})}{\sum_{j \neq i} ((1 - t) \cdot w_{ij} + t \cdot z_{ij})},$$  

$$\dot{y}_i \leftarrow \frac{\sum_{j \neq i} ((1 - t) \cdot w_{ij} + t \cdot z_{ij}) \cdot (y_j + d_{ij} \cdot (y_i - y_j) \cdot b_{ij})}{\sum_{j \neq i} ((1 - t) \cdot w_{ij} + t \cdot z_{ij})}.$$  

In the majorization, the radial constraints are not directly and immediately enforced; rather, the main visual features of the initial configuration are preserved. Then the bias is shifted from the distance component towards the radial component by gradually increasing $t$ from 0 to 1. When the number of iteration steps
k is predefined, a linear interpolation gives values \( t = 0, \frac{1}{k}, \frac{2}{k}, \ldots, \frac{k-1}{k}, 1 \); otherwise, the iterative process may be simply be repeated with a sequence of values for \( k \) converging to 1 from below until the layout is sufficiently stable. Using either variant, in each step, a slightly different objective function is sought to be minimized, and the current iterate preconditions the next step, thus keeping the series of iterates continuous.

In the multidimensional scaling literature, occasionally a distinction is made between weakly and strongly constrained MDS problems [13]. In the former, the solutions are allowed to deviate from the given constraints, and this deviation is penalized by additional stress; in the latter, only solutions which exactly satisfy the constraints are feasible. In a way, a strongly constrained MDS problem can be thought of as a special case of a weakly constrained problem, in which the deviation penalty is zero.

In this terminology, setting \( t = 1 \) in (8) turns a weakly constrained problem into a strongly constrained one, provided that the set of constraints is realizable, i.e., a solution with zero constraint stress exists. In all other cases, it should be noted that, even though the distance component vanishes when \( t \to 1 \), minimizing \( \sigma_t(p) \) is not the same as minimizing \( \sigma_Z(p) \) because of the running preconditioning described above.

### 4 Radial Layout

#### 4.1 Target Diagrams

The focus is put on a node by emphasizing the visual display of its vicinity, constraining all others to attain Euclidean distances corresponding to their graph-theoretical distances, i.e., relative to the focused node, all structural distance-\( k \) neighborhoods are mapped to a geometric \( k \)-neighborhood.

The constraint weight matrix takes only pairs of nodes into account in which the focused node is involved, while reducing all other weights to zero. Without loss of generality, let \( n \) be the index of the node to be focused. \( D \) and \( W \) are defined as above, and the constraint weight matrix \( Z = (z_{ij}) \) has only zero entries except for the \( n \)-th row and column, which contains weights

\[
z_{ij} = \begin{cases} w_{ij} & \text{if } i = n \text{ or } j = n \\ 0 & \text{otherwise} \end{cases}
\]

 derived from the distances to the focal node, so that interpolating from \( W \) to \( Z \) gradually increases the focal node’s relative impact on the configuration.

A famous social network was studied by Zachary and, subsequently, many other sociologists. It describes the friendship relations among 34 members in a karate club in a US university in the 1970s [21]. Over the course of a two-year study, the network breaks apart into two clubs because of disagreements between the administrator and the instructor, who leaves the club and takes about half of the members with him. Fig. 2 shows how the same initial layout is modified to a radial layout focused on the instructor (a) and the administrator (b).
4.2 Centrality Drawings

When the radial constraints do not directly correspond to one of the columns in the distance matrix, we assume that this additional input is given as a vector \( r = [r_1, \ldots, r_n]^T \in \mathbb{R}^n \), with \( r_i \geq 0 \) for all \( i \in \{1, \ldots, n\} \). The radial constraints can be formulated in terms distances to the center, added to the distance matrix; since node \( i \) is located on a circle with radius \( r_i \) if its Euclidean distance to the center is equal to \( r_i \), and the center has coordinates \((0, 0)\), this is equivalent to

\[
\|p(i)\| \approx r_i. \tag{12}
\]

The origin is treated as an additional dummy node indexed with \( n + 1 \). The stress majorization procedure is applied to a layout problem of \( n + 1 \) objects; in [3] such a dummy is used to enforce a circular configuration by using the same radius for all objects. The distance and weight matrices involved in (8) are

\[
D = \begin{bmatrix}
  d_{11} & \cdots & d_{1n} & r_1 \\
  \vdots & \ddots & \vdots & \vdots \\
  d_{n1} & \cdots & d_{nn} & r_n \\
  r_1 & \cdots & r_n & 0
\end{bmatrix},
\]

\[
W = \begin{bmatrix}
  d_{11}^{-2} & \cdots & d_{1n}^{-2} & 0 \\
  \vdots & \ddots & \vdots & \vdots \\
  d_{n1}^{-2} & \cdots & d_{nn}^{-2} & 0 \\
  0 & \cdots & 0 & r_i^{-2}
\end{bmatrix},
\]

\[
Z = \begin{bmatrix}
  0 & \cdots & 0 & r_i^{-2} \\
  \vdots & \ddots & \vdots & \vdots \\
  \vdots & \ddots & \vdots & \vdots \\
  0 & \cdots & 0 & r_n^{-2}
\end{bmatrix}. \tag{13}
\]

**Fig. 2.** Radial layouts of Zachary’s karate club network \((n = 34, m = 77)\), by weight interpolation, for \( t \in \{0, 0.9, 1\} \). Members leaving with the instructor are shown as yellow squares, members staying with the administrator as red circles.
Let $c = [c_1, \ldots, c_n]^T$ be a centrality measure on the nodes of graph $G$. For every node $i \in V$ its radius is given by

$$r_i = \frac{\text{diam}(G)}{2} \cdot \left(1 - \frac{c_i - \min_{j \in V} c_j}{\max_{j \in V} c_j - \min_{j \in V} c_j + c(G)}\right),$$

where multiplying with half the diameter serves to keep distances and the radial constraints on the same scale, and $c(G)$ is an offset parameter that prevents more than one maximally central nodes from coinciding in the center [4]. Simplified pseudo-code, which is targeted at radial constraints, is given in Algorithm 1. The majorization is realized as a local variant, in which the coordinates are updated node-by-node immediately.

For dynamic visualization scenarios, an inherently smooth transition between layouts with different foci is obtained by simply using the intermediate layouts given by the steps in the majorization process. In the transition from one focus to the other, it is advantageous to not directly interpolate between the two corresponding constraint weight matrices, but to take a detour via the original weight matrix having entries $d^{-2}_{ij}$, so as to re-introduce all the shortest-path distances to remove artifacts potentially introduced after focusing on the first node.

4.3 Travel Time Maps

When traveling with a public transportation system, schematic maps are essential for many users. Such maps depict lines, stations, zones, and connections to other traffic systems. Since the primary use for such a map is travel planning within this network, usability and readability are more important criteria than
Algorithm 1. Radial layout

**Input**: connected undirected graph $G = ((1, \ldots, n), E)$, radii $r_1, \ldots, r_n \in \mathbb{R}^+$

**Output**: coordinates $x, y \in \mathbb{R}^n$ with $\sqrt{x_i^2 + y_i^2} = r_i$ for all $i \in \{1, \ldots, n\}$

$D \leftarrow$ matrix of shortest path distances $d_{ij}$

$W \leftarrow$ matrix of inverse squared distances $d_{ij}^{-2}$

$(x, y) \leftarrow$ layout with low $\sigma_W(x, y)$

for $t = 0, \frac{1}{2}, \frac{3}{2}, \ldots, 1$ do

for all $b_{ij}$ as in (5)

foreach $i \in \{1, \ldots, n\}$ do

$x_i \leftarrow \sum_{j \neq i} (1 - t) \cdot w_{ij} (x_j + d_{ij} \cdot (x_i - x_j) \cdot b_{ij}) + t \cdot r_i^{-2} (x_i, a_i)$

$y_i \leftarrow \sum_{j \neq i} (1 - t) \cdot w_{ij} (y_j + d_{ij} \cdot (y_i - y_j) \cdot b_{ij}) + t \cdot r_i^{-2} (y_i, a_i)$

end foreach

end for

end for

the accurate representation of actual geographic positions. In the graph drawing literature, this drawing style is called metro map layout [14].

One of the most prominent schematic maps is Beck’s famous London tube map; it has been and is still being reworked and improved and has inspired similar maps for systems of public transportation all over the world. While schematic maps are widely perceived as very useful, a potential drawback is that they tend to distort a user’s perception of closeness, thus compromising the decisions made in the travel planning process, e.g., because stations are displayed as more proximate than they actually are.

If the starting and ending stations of a planned journey are known, the radial constraints can be used to highlight the time needed for traveling between them, by focusing only on one, as described above. In addition, shortest paths between the two stations can be highlighted by putting the focus on both of them at the same time.

Let the nodes in the focus be $n-1$ and $n$, without loss of generality. Again, $D, W \in \mathbb{R}^{n \times n}$ are defined as the matrices of shortest-path distances and their inverse squares, respectively. The weight matrix used is $Z = (z_{ij})_{ij} \in \mathbb{R}^{n \times n}$ with

$$z_{ij} = \begin{cases} w_{ij} & \text{if } i \in \{n-1, n\} \text{ or } j \in \{n-1, n\} \\ 0 & \text{otherwise} \end{cases}$$

and contains a $(n - 2) \times (n - 2)$ submatrix with zero entries.

We use the connection graph of the London tube with estimated minimal travel times obtained from the Transport for London web site, or derived from the geographic distance, where estimates are not available. For the sake of simplicity, we did not consider the time needed to walk from one track to the other when changing lines (see also [6]).
Figure 4. Radial layouts of the London Tube graph using estimated travel times. The concentric circles indicate travel times in multiples of 10 minutes. The stations are constrained to be at distance equal to their minimum travel times.

Radial layouts are given in Fig. 4, where stations are placed at a distance from the center proportional to their estimated minimum travel times from (a) Golders Green and (b) Greenwich independently and (c) from both stations at the same time.

5 Conclusion

Radial constraints fit well into the framework of multidimensional scaling by stress majorization because the radii can be expressed in terms of desired
Euclidean distances, which requires only minor modifications of available implementations of the stress minimization.

Motivated by the results obtained from the relatively simple radial constraints, and other experiments, we feel that they deserve more attention, because they allow the aesthetic goals of the visualization results to be explicitly formulated and quantified, and can be easily plugged into existing algorithms.

We think that, with the careful choice of a weighting scheme, the ideas presented above are easily carried over to layout tasks with more general constraints, such as the display of grouping, the computation of dynamic layouts, and the visualization of edge strength, certainty, or probability.

References

   http://www.tom-carden.co.uk/p5/tube_map_travel_times/applet/


18. Northway, M.L.: A method for depicting social relationships obtained by sociometric testing. Sociometrics 3, 144–150 (1940)

