Stott, Jonathan, Rodgers, Peter, Martinez-Ovando, Juan Carlos and Walker, Stephen G. (2011) ***Automatic Metro Map Layout Using Multicriteria Optimization.*** **Transactions on Visualization and Computer Graphics, 16 (1). pp. 182-196.**

Automatic Metro Map Layout Using Multicriteria Optimization

Jonathan Stott, Peter Rodgers, Juan Carlos Martínez-Ovando, Stephen G. Walker
University of Kent, UK

Abstract—This paper describes an automatic mechanism for drawing metro maps. We apply multicriteria optimization to find effective placement of stations with a good line layout and to label the map unambiguously. A number of metrics are defined, which are used in a weighted sum to find a fitness value for a layout of the map. A hill climbing optimizer is used to reduce the fitness value, and find improved map layouts. To avoid local minima, we apply clustering techniques to the map — the hill climber moves both stations and clusters when finding improved layouts.
We show the method applied to a number of metro maps, and describe an empirical study that provides some quantitative evidence that automatically-drawn metro maps can help users to find routes more efficiently than either published maps or undistorted maps. Moreover, we found that, in these cases, study subjects indicate a preference for automatically-drawn maps over the alternatives.

Index Terms—information visualization, diagram layout, graph drawing.

1 Introduction

Since Harry Beck developed the iconic map of the London Underground, first published in 1933 [1], [2] similar schematic diagrams have been used to guide travellers on public transport networks. Typically, such diagrams are produced by modifying the network layout so that unnecessary complexity is removed. For example, lines run at regular angles, stations are evenly spaced and labels are placed in unambiguous locations. Whilst the geometry of the map is changed, the topology is retained. The great advantage of such diagrams over undistorted maps is that they simplify the key tasks of route planning and navigation for travellers.

Currently, schematic diagrams are produced by human designers and take a considerable time to generate. Whilst this may be sensible for the use of such maps as static views of entire networks, there are other applications for which automated layout would be a great benefit. Cheap and quickly produced computer generated schematic maps might be used for personal travel plans, generating diagrams for networks not currently included in current maps, or for using schematic networks in other application areas, such as water or gas utility networks [3]. In addition, the metro map metaphor, which draws graph-based data in a similar form to metro maps, has been widely used in non-geographical application areas, where the lack of geographical constraints allow more freedom in diagram layout. Such applications include project plans [4], cancer pathways [5] and web site mapping [6].

Our paper describes a mechanism for drawing usable metro maps. This is achieved with a new method for metro map layout, multicriteria optimization, which performs the difficult task of generating a good line layout with unambiguous, readable labels. A number of metrics are defined, which are used in a weighted sum to attempt to measure the aesthetic quality of the diagram. Our approach also uses three new clustering mechanisms to avoid local minima. The usability of metro maps produced by our system has been tested by empirical study and we describe the experiments and statistical analysis that brings us to the conclusion that our metro map layout method can produce usable diagrams. Previous conference publications [7], [8] have described early versions of the method that appears in Section 3. However, the work in this section is extended from the early publications, with improved criteria and optimization. The clustering and empirical research is new work and has not been previously published. The work given here formed the basis of a Ph.D. thesis [9].

The rest of this paper is organized as follows: Section 2 describes some background in the layout of metro maps and other schematic diagrams; Section 3 describes our optimization method; Section 4 describes how the optimizer is extended with clustering methods; Section 5 gives some examples of metro maps drawn with the system; Section 6 describes the empirical study and gives our interpretation of results; finally Section 7 gives our conclusions and suggests further work.

2 Related Work

Other approaches to metro map layout have not yet been empirically tested for their effectiveness in drawing real-world examples of metro maps. These methods, in the broad, attempt to achieve similar layouts to those we attempt, with similar diagram criteria either implicitly or explicitly specified. However, unlike our method, labelling is typically not attempted at all or performed after the diagram layout has completed.

The first attempt at automatic metro map layout was by Hong *et al.* who use a force-directed approach to laying out metro maps [10], [11]. They use combinations of different forced-based algorithms applied sequentially. Labelling occurs after the diagram layout, but because this does not allow the map layout to make room for labels, occlusion and ambiguity can still occur. Other problems exist because they do not consider the geography of the map, resulting in counter intuitive layouts, for example, stations geographically to the north of others can be placed to the south. In practice, the implicit definition of criteria through the force method tends to result in irregular spacing of stations and non-discrete line angles typically appear, which are not a feature of most published schematic metro maps.

Nöllenburg and Wolff describe a method of drawing metro maps using mixed-integer linear programming [12] which extends linear programming by introducing the notion of constraining variables to be within certain discrete integer ranges. Constraints include octilinear lines (horizontal, vertical or $45°$ diagonal), maintaining a minimum line length, minimum line bends and minimum total line length. The final diagrams lack labels. Note that both the work of Nöllenburg and Wolff and that of Hong, outlined in the previous paragraph, use variants of the Sydney metro map as examples, allowing for comparison with the Sydney map generated by the system described in this paper, Figure 22.

Merrick and Gudmundsson describe a path simplification method which restricts the number of directions that lines can take [13]. They simplify the lines in order of importance, determined using a heuristic function based on the number of interchange stations on the line. The method fails to maintain topology, does not achieve an effective overall structure and lacks labelling, however it produces results in a relatively quick time. Other research includes efforts by Bekos et. al. to minimize line crossings for embedded metro maps [14], which they achieve for restricted types of diagrams.

In addition, there are similar generalization and schematic problems for application areas other than metro maps. Agrawala and Stolte describe a simulated annealing system for producing simplified route maps [15]. This work uses criteria of: length generalization to even out route segments; angle generalization to prevent small turning angles; and shape generalization to simplify the shape of roads. Avelar and Huber [16] show a similar method but model their route maps on the characteristics of public transport networks.

Casakin *et al.* [17] provide a taxonomy of various aspects of schematic route maps (particularly intersections), and use their taxonomy to provide an empirical assessment of schematic graphs. Yates and Humphreys [18] give a discussion of various aspects of schematic diagrams and show a prototype (which uses a heuristic provided as a sample applet in the Java 1.1.6 SDK). Cabello *et al.* [19] presents a relatively efficient combinatorial algorithm for the generation of schematic maps which takes into account a number of requirements such as choosing the minimum separation of stations and not moving stations. Lauther and Stübinger [20] present a demonstration of software which is capable of laying out schematic diagrams using a force-directed approach with the aim of visualizing cable plans schematically.

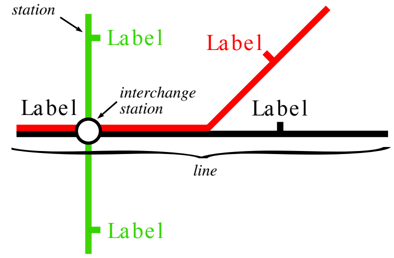


Fig. 1. Metro map features

The methods for route map schematic layout are generally successful for the problems that they try to solve. However, the size of diagram is generally smaller and less complex than would be expected for a metro map.

3 Optimization Method

This section describes the basic concepts of our metro map layout method. We detail the hill climbing optimizer, the criteria measured, and method for combining criteria to produce a fitness value. This section only discusses individual station movement; Section 4 discusses the movement of clusters of stations.

A *diagram*, G is a set of *stations*, V, with connections between pairs of stations represented by a set of edges, E. When drawing metro maps, we use an *edge* to represent a single connection between two stations. In some cases, there may be several edges connecting two stations where two or more metro lines run together. We use the term *metro line* to represent a subset of edges that form a particular line on the network (such as the Central or Northern Lines on the London Underground map). Edges also have metadata in the form of a colour that identifies which line they are part of. These features are illustrated in Figure 1.

The diagram is embedded on an integer grid, meaning that stations must be centred on grid intersections, however, edges do not have to follow grid lines. The spacing between adjacent intersections in the grid is denoted by g and is always large enough to allow parallel edges between stations to be placed without ambiguity. Making the search space discrete in this manner allows us to dramatically reduce the number of potential locations for stations. We also define a preferred multiple of the grid spacing for station separation, l, making the ideal edge length lg.

The method has been tested on nine real world maps to date, which can be seen in [9], in addition to a number of diagrams constructed to test particular issues. See also the Appendices of this paper for examples. The number of real world test maps is restricted by the difficulty of getting data, as each undistorted map must be encoded

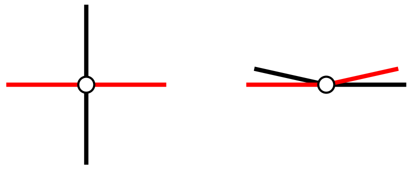


Fig. 2. Examples of optimal angular resolution (left) and poor angular resolution (right).

by hand. The criteria we selected includes criteria that empirical research suggests are effective in the field of graph layout. Not all such criteria are appropriate for metro map layout (such as symmetry). In addition, the line straightness and balanced edge length criteria have been added for the particular requirements of metro map layout. The clustering mechanisms were based on informal examination of the output whilst the system was being developed.

3.1 Hill Climbing

Multicriteria optimization has been used previously in graph drawing [21], [22]. These previous methods use genetic algorithms or simulated annealing to optimize a fitness function, however we found that a simpler method using hill climbing was more appropriate for this application. Simulated annealing adds an element of non-determinism in order to escape from local minima in the search space, but a larger number of iterations would be necessary to reach a minimum in the search space. Moreover, the local minima that typically occur in schematic networks are better dealt with by clustering. Genetic algorithms are also non-deterministic, and converge more slowly than hill climbers or simulated annealers.

In outline, our method operates in this manner: firstly we find an initial layout for the stations, which is the undistorted layout, but with the stations snapped to the grid and only one station at any point. For each station in the diagram we calculate the fitness of the diagram. We then search the points around a rectangle centred on the station at a given distance. We then move the station to the location that most improves the fitness. If none do, then the station is not moved. This is performed for all stations in the diagram. We then see if the label layout in the diagram can be improved in a similar manner. Once all station positions and label positions have been tested, the process is repeated until no more improvement can be made. On each of these repetitions, a cooling factor reduces the search rectangle to allow fine tuning of layout as the search progresses. The process in detail, including the clustering step, discussed in Section 4, is given in Algorithm 1.

3.2 Station Criteria

Movement of stations depends on the calculation of the weighted sum of several criteria which are judged to

Algorithm 1 Metro Map Layout

$G \Leftarrow (V, E, L)$
snapStations(V)
$m_{T0} \Leftarrow$ calcStationCriteria(V) + calcLabelCriteria(L)
$running \Leftarrow$ **true**
while $running$ **do**
  for $v \in V$ **do**
    $m_{N0} \Leftarrow$ calcStationCriteria(V)
    $m_N \Leftarrow$ findLowestStationCriteria(V)
    if $m_N < m_{N0}$ **then**
      moveStation(v)
    end if
  end for
  $P \Leftarrow$ clusterOverlengthEdges$(V, E) \cup$ clusterBends$(V, E) \cup$ clusterPartitions(V, E)
  for $p \in P$ **do**
    $m_{N0} \Leftarrow$ calcStationCriteria(V)
    $m_N \Leftarrow$ findLowestStationCriteria(V)
    if $m_N < m_{N0}$ **then**
      moveCluster(p)
    end if
  end for
  for $l \in L$ **do**
    $m_{L0} \Leftarrow$ calcLabelCriteria(L)
    $m_L \Leftarrow$ findLowestLabelCriteria(L)
    if $m_L < m_{L0}$ **then**
      moveLabel(l)
    end if
  end for
  $m_T \Leftarrow$ calcStationCriteria(V) + calcLabelCriteria(L)
  if $m_T \not< m_{T0}$ **then**
    $running \Leftarrow$ **false**
  else
    $m_{T0} \Leftarrow m_T$
  end if
end while

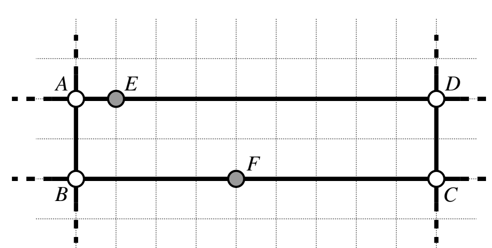


Fig. 3. Balanced edge lengths.

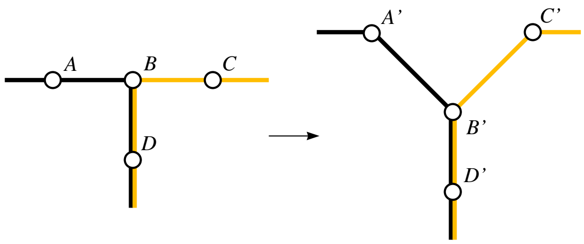


Fig. 4. Examples of poor line straightness (left) and improved line straightness (right).

affect the aesthetic quality of the map. Our basis for the selection of criteria comes from existing research that evaluates aesthetic criteria in relation to graph drawing [23] as well as criteria considered specific to the aesthetics of schematic diagrams and metro maps [10], [17], [24], [25]. The criteria evaluate to a lower value when improved. The station criteria are:

- **Angular Resolution Criterion,** c_{N1}**.** The angles of incident edges at each station should be maximized, because if there is only a small angle between any two adjacent edges then it can become difficult to distinguish between them. See Figure 2. It is calculated by

$$c_{N1} = \sum_{v \in V} \sum_{\{e_1, e_2\} \in E_v} \left| \frac{2\pi}{\rho(v)} - \theta(e_1, e_2) \right| \tag{1}$$

  where $\rho(v)$ is the degree of the station v (the degree of a station is the count of its incident edges) and $\theta(e_1, e_2)$ is the angle in radians between two adjacent edges e_1 and e_2 incident to v.
- **Edge Length Criterion,** c_{N2}**.** The edge lengths across the whole map should be approximately equal to ensure regular spacing between stations. It is based on the preferred multiple, l, of the grid spacing g. The purpose of the criterion is to penalize edges that are longer than or shorter than lg. It is calculated by

$$c_{N2} = \sum_{e \in E} \left| \frac{|e|}{lg} - 1 \right| \tag{2}$$

  where $|e|$ is the length of edge e.
- **Balanced Edge Length Criterion,** c_{N3}**.** The length of edges incident to a particular station should be similar. One of the characteristics of metro maps is that there are many stations with two incident edges (degree two). Figure 3 shows an example whereby there are two stations, E and F, with degree two. If we are only considering the edge length criterion for these two stations, it evaluates to the same value for both stations. However, we want to ensure that the edge lengths are similar. In these cases, the balanced edge length criterion can help by penalizing stations with degree two that have incident edges with unbalanced lengths. It is calculated as the sum of the absolute difference between the lengths of the two incident edges of every degree two station in the diagram

$$c_{N3} = \sum_{v \in V, \rho(v)=2} ||e_1| - |e_2|| \tag{3}$$

  where e_1 and e_2 are the incident edges of station v which has degree $\rho(v) = 2$.
- **Line Straightness Criterion,** c_{N4}**.** Edges that form part of a line should, where possible, be collinear either side of each station that the line passes through. One of the important features of metro maps is that metro lines appear to pass through stations so that the entry edge is more-or-less directly opposite the exit edge. This is particularly important if there are two or more lines passing through a station, see Figure 4. It is calculated by

$$c_{N4} = \sum_{(v \in V)} \left(\sum_{e_1, e_2 \in E} \theta\left(e_1, e_2\right) \right) \tag{4}$$

  where $\theta\left(e_1, e_2\right)$ is the smallest angle between adjacent edges e_1 and e_2, and e_1 and e_2 are the only two edges of the same line that are incident to the station v.
- **Octilinearity Criterion,** c_{N5}**.** Each edge should be drawn horizontally, vertically or diagonally at 45°, so we penalize edges that are not at a desired angle. It is calculated by

$$c_{N5} = \sum_{\{u,v\} \in E} \left| \sin 4 \left(\tan^{-1} \frac{|y(u) - y(v)|}{|x(u) - x(v)|} \right) \right| \tag{5}$$

  where $\{u, v\}$ is an edge between stations u and v, and $y(v)$ and $x(v)$ are the y- and x-coordinate of station v respectively.

Figure 5 shows an example that illustrates the octilinearity criterion. The result of calculating the criterion for each edge in the example graph is shown in Table 1. As is expected, edges which are already at an angle of some multiple of 45° (AB and FG) evaluate to zero, whereas edges which are at angles furthest from multiples of 45° evaluate to the highest values. Edges BC and BF evaluate to the same value because they are both 18.43° away from the nearest multiple of 45°.

TABLE 1
Examples of octilinearity criterion calculations.

Edge, $e = \{u, v\}$	c^{e}_{N5}
$\{A, B\}$	$\left\|\sin 4 \left(\tan^{-1} \frac{0}{4}\right)\right\| = 0$
$\{B, C\}$	$\left\|\sin 4 \left(\tan^{-1} \frac{2}{6}\right)\right\| = 0.96$
$\{C, D\}$	$\left\|\sin 4 \left(\tan^{-1} \frac{3.8}{0.6}\right)\right\| = 0.586$
$\{B, E\}$	$\left\|\sin 4 \left(\tan^{-1} \frac{6}{0.6}\right)\right\| = 0.388$
$\{B, F\}$	$\left\|\sin 4 \left(\tan^{-1} \frac{1}{2}\right)\right\| = 0.96$
$\{F, G\}$	$\left\|\sin 4 \left(\tan^{-1} \frac{0}{4}\right)\right\| = 0$

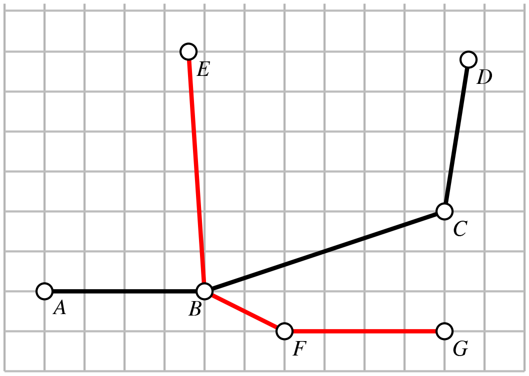


Fig. 5. Octilinearity example.

Fig. 6. Example of the enforcement of the relative positions when moving a station. The grey shaded area shows the degree of freedom afforded to station A (left) and station C (right).

3.3 Station Rules

As well as the above five criteria, we have implemented four station movement rules which are strictly enforced during the layout process. We apply rules in addition to criteria to enforce particularly important features of a schematic diagram. The four station movement rules are:

- **Bounding Area Restriction Rule.** Restrict the movement of stations to be within a certain bounding area so that the final diagram will fit on the target display. When multiple stations are moved, no movement may place any of the stations beyond the bounding area.
- **Relative Position Rule.** Enforce the relative position between adjacent of stations. Although metro maps are a generalization of the undistorted network, relationships such as one station being north of another station are still important to the usability of the drawn map. This rule ensures that the relative positions between neighbouring stations do not change. Figure 6 illustrates the effect of enforcing relative positions. Stations may move only within the quadrant in which they start. There are four possible quadrants, defined by the division of the plane by two orthogonal axes centred on the relevant neighbouring station. If the station to be moved starts on the border of two quadrants (because it is horizontally or vertically aligned with the neighbour), then it may move in either quadrant.
- **Occlusions Rule.** Avoid the introduction of occlusions of other edges and stations to ensure that a station is not moved so that it is not lying on top of any other station or edge, and that edges do not cross other edges or lie on top of any other station.
- **Edge Ordering Rule.** Preserve the ordering of edges incident to a station. The relative positions rule allows us to restrict the relative positions between two stations. However, there are limitations to this rule that mean that the topology of the diagram could be changed by the movement of a station, see Figure 7. To implement this rule we need to find the clockwise ordering of edges around the station being moved and any neighbouring station in the diagram. At each potential new location for a station, the edge ordering is checked and the location disregarded if the orderings change.

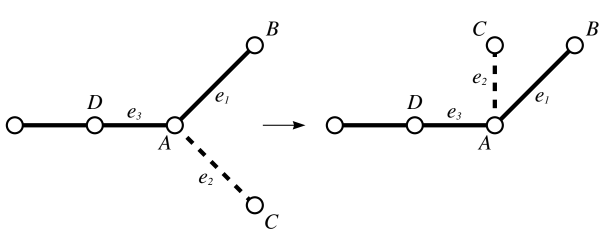


Fig. 7. Preservation of edge ordering. Without preserving the ordering of edges, station C would be able to move as shown, changing the topology of the map.

3.4 Station Movement Criteria Weightings

Each criterion has an independent weighting. The intention of these weightings is twofold. First, the functions generate values which can vary by an order of magnitude or more between each of the criteria. The weightings allow the values of each criterion to be brought within the same magnitude of each other (normalized). It is not possible to bound all criteria to upper and lower values, then scale to between 0 and 1, because many of the station criteria are unbounded. However, it is still important to ensure that one criterion does not completely overwhelm the other criteria. Second, by using a higher weighting, a preference can be placed on a particular criterion if the effects of that criterion are required to be more prominent. Conversely, a lower weighting can be used to reduce the effect of a particular criterion, down to zero if it is not appropriate for that particular case.

The sum of the weighted criteria for station movement, m_N, is given by

$$m_N = \sum_{i=1}^{5} w_{Ni} c_{Ni}. \tag{6}$$

where w_{Ni} is the weighting for criterion c_{Ni}. The weightings that we used were determined through a process of trial and error. This process first involved setting the weightings such that the weighted values are effectively normalized (to cancel out differences in magnitudes) and then using particular examples to determine how each

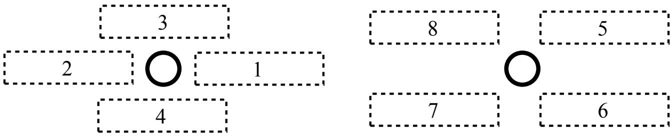


Fig. 8. Search space for labelling the metro map.

TABLE 2
Label position values. The positions refer to the positions shown in Figure 8.

	position	value
1	east	1.0 (best)
2	west	1.1
3	north	1.4
4	south	1.4
5	north-east	1.5
6	south-east	1.6
7	south-west	1.7
8	north-west	1.8 (worst)

weighting should be modified so that it has the desired effect. The process was driven by the subjective judgement of the investigators, as a more formal approach would have taken more time than was available.

Appendix A shows the effect of removing each station movement criteria in turn, indicating that all criteria have some effect on the final result.

3.5 Label Criteria

Labelling is an integral part of metro maps — in an informal discussion with a professional metro map designer, we discovered that labelling was considered the major issue in the layout of the London Underground map. Hence, it should form an integral part of any schematics layout method. Similarly with station and line optimization, we have designed a number of criteria for label placement. These criteria are based on cartographic point labelling considerations [26]. However, it should be noted that some principles differ in metro map layout, in particular, positions directly to the left and right of the station are acceptable in metro map layout because the line prevents the station being misinterpreted as a type character in the label. The advantage of calculating a fitness function for labelling, rather than applying alternative, more widely used methods, is that the labelling can be integrated with the station layout in the hill climber.

In order to reduce the number of potential locations for labels and to allow a preference for one position over another, we limit the number of positions using a labelling space. Figure 8 shows our chosen labelling space, which allows eight different label positions. The values for the positions are currently independent of line orientation, however, they could be adjusted for particular line orientations for greater flexibility.

Occasionally a label might contain a large amount of text with several words so we split a long label length of the label if it exceeds $0.75lg$. This is a relatively simplistic

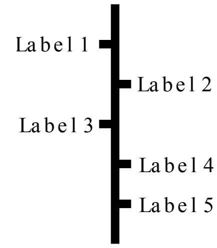


Fig. 9. Label position consistency.

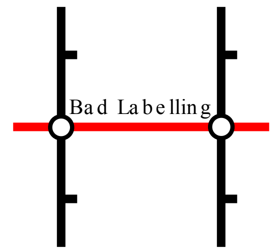


Fig. 10. An example of ambiguous labelling.

strategy, as whilst this usually ensures the label fits in between stations, often a split is made when there is still plenty of room for it in the diagram.

The seven labelling criteria are:

- **Label Occlusion Criteria,** c_{L1}, c_{L2}, c_{L3}. These three criteria count the number of stations, edges and other labels that intersect/occlude labels respectively. As intersections drastically reduce the readability of the map, it is highly desirable to ensure that they happen as infrequently as possible. However, there may be occasions where the readability of the diagram would be improved if a label were allowed to occlude an edge, as in dense areas of the diagram it may not be possible to find any improvements to the position of labels to resolve all label occlusions.
- **Label Position Criterion,** c_{L4}. Places a preference on label positions in the labelling space by putting a value on each position. A label can occupy any one of the eight locations in the labelling space shown in Figure 8. Some label positions are more preferential than others, so each different position in the labelling space is assigned a value relating to the preference for that position. Table 2 shows

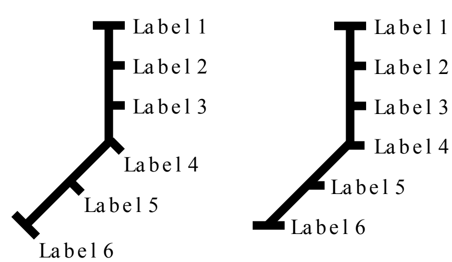


Fig. 11. Examples of perpendicular tick labels (left) and non-perpendicular tick labels (right)

the set of values for each position in the labelling space. The label position criterion is then defined as the sum of the position values for each label in the diagram.

- **Label Position Consistency Criterion,** c_{L5}**.** Gives preference to labels along a line in the map that consistently appear on the same side of the line. This improves readability because the labels appear as a list which can be read easily rather than having to switch attention from one side of the line to the other. The criterion is only calculated for labels with one or two neighbouring stations, as stations with more than two neighbours have to be consistent with more than one line. The calculation is fairly simple: for each station in the diagram with degree equal to two, a count is kept of the number of times the position of the label of an adjacent station (if that station has degree less than three) differs to the position of the current station, counting each pair of stations once. Figure 9 shows an example of poor label position consistency where the unweighted value of the label position criterion would be three.
- **Station Proximity Criterion,** c_{L6}**.** Penalizes labels that come into close proximity to unrelated stations, so discouraging labels from being positioned too close to other unrelated stations, which causes ambiguity when deciding which station the label relates to. It is calculated by

$$c_{L6} = \sum_{k \in L} \sum_{v \in V, k_v \neq k} \frac{1}{d\left(k, v\right)^2} \tag{7}$$

where $d\left(k, v\right)$ is a function giving the distance from the closest point on the bounding box of label k to station v. Notice that we are interested in $n \in V, k_v \neq k$, that is all stations in the diagram except the one for which the label (k_v) is the label we are considering (k). In other words, we do not take into account the distance between a label and station that that label belongs to. In practice, because most stations in the diagram will be some way from the label in question, they will contribute very little to c_{L6}. We can therefore approximate the contribution of stations with $d\left(k, v\right) > x$ to zero. We use a value of g for x.
- **Perpendicular Tick Criterion,** c_{L7}**.** Encourages the tick (and therefore the position of the label) for a particular station to be perpendicular to the line. Figure 11 illustrates the need for this criterion. The left-hand diagram shows a line where the ticks showing stations have been drawn perpendicular to the line. The right-hand diagram shows ticks always drawn straight to the right (labels are positioned to the east). While the labels and ticks for the vertical part of the line remain the same, the perpendicular ticks on the diagonal part of the line are more prominent. The minimum distance between the line and the labels on the diagonal part is also greater when the labels are drawn diagonally, but the association with the relevant tick is not lost. For a station v, $e1_v$ and $e2_v$ are the connecting edges and $\theta 1_v$ and $\theta 2_v$ are the angles between the tick and $e1_v$ and $e2_v$ respectively. The unweighted value of this criterion for a single station is the absolute difference between the two angles. The total value for all stations, V, in the graph is therefore

$$c_{L7} = \sum_{v \in V} \left| \theta 1_v - \theta 2_v \right| . \tag{8}$$

3.6 Labelling Criteria Weightings

The values for label weightings were determined through trial and error with various examples in a similar manner to the way that we determined station movement criteria weightings, see Section 3.4, and as with the station criteria, the label criteria are not restricted to between 0 and 1, for consistency. However, unlike the station criteria, the label criteria can be bound by dividing by the number of stations in the diagram.

The sum of the weighted criteria for labelling, m_L, is given by

$$m_L = \sum_{i=1}^{7} w_{Li} c_{Li}. \tag{9}$$

where w_{Li} is the weighting for criterion c_{Li}. As with the station movement criteria weightings, the values for the label weightings can be modified by the user depending on the characteristics of the particular metro map being drawn.

4 Clustering

Section 3.1 introduced a method for laying out metro maps using multicriteria optimization that improved the fitness in the diagram by moving individual stations. This often results in easily identifiable cases of local minima that seem improvable if clustered groups of stations were moved together.

We have three methods for clustering:

- clustering based on overlength (or underlength) edges;
- clustering based on bends in lines;
- clustering based on partitioning the diagram into two parts that can be moved closer together.

Once clusters have been identified, they are moved in exactly the same way that individual stations are moved with the only difference being that all the stations in the cluster are moved and the relative position of stations in a cluster is maintained.

As with station movement criteria, the effect of removing each clustering method in turn is shown in Appendix A, indicating that all the clustering methods have some effect on the final result.

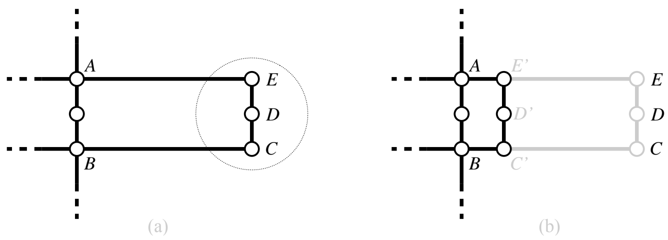


Fig. 12. Clustering multiple overlength edges. The edges AE and BC are too long (a) and it is only possible to reduce the length of these edges by moving stations C, D and E at the same time (b).

4.1 Clustering Overlength Edges

A frequent problem we encountered when experimenting with our layout system was that of long edges that do not reduce in length. We define overlength edges as being edges which are longer than lg (the ideal edge length). Figure 12 shows such an example with two overlength edges. If we only allow one station to move at a time, the overlength edges connecting the two groups of stations cannot reduce in length.

Our first attempt at solving this problem, given in [8] attempted to cluster groups of stations at the ends of lines. However this does not always deal with multiple overlength edges, as shown in Figure 12. Instead, we cluster the diagram into groups of stations connected by ideal length edges.

4.2 Clustering Non-Straight Lines

Many lines contain short deviations or kinks. This occurs when fitting a slightly off-straight line to the grid or where three stations are too close together to fit onto the grid without the middle station being offset relative to the rest of the line.

To improve non-straight lines we identify clusters of stations by looking at stations which have exactly one or two neighbours. Figure 13 shows an example of clusters found with this method. This means that stations A and F are discounted from forming part of a cluster from the outset (and could even be removed from the graph while we are searching for clusters). Clusters are then identified by finding the minimum set of connected stations which are collinear.

4.3 Partitioning

The results of experiments on test maps also identified local minima that occur because overlength edges cannot always be reduced by the clustering as described in Section 4.1. An example is shown in Figure 14. However, improvements can be made by partitioning the diagram into two along overlength edges and treating these partitions as clusters.

Our approach to finding partitions in the graph can be summarized as follows:

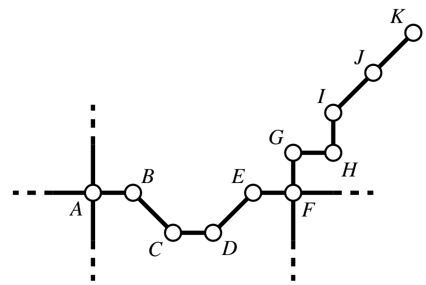


Fig. 13. Clustering stations to find non-straight lines. Ultimately, six clusters will be identified in this graph: $\{BC\}$, $\{CD\}$, $\{DE\}$, $\{GH\}$, $\{HI\}$ and $\{IJK\}$.

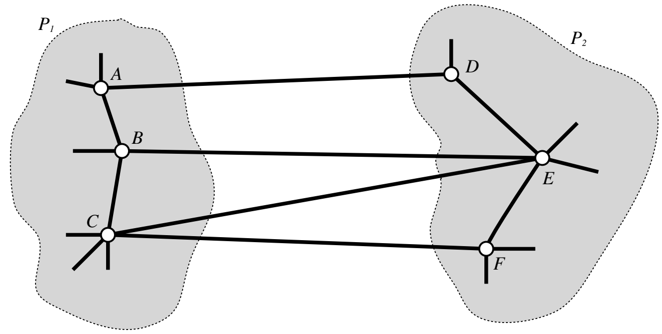
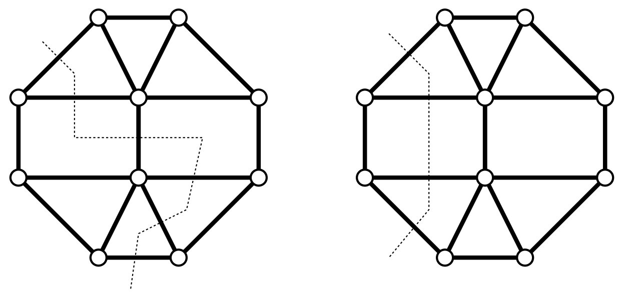


Fig. 14. Partitioning a diagram into two. The edges AD, BE, CE and CF are all too long but the only way of shortening them is to move either the left-hand partition P_1 or the right-hand partition P_2. Both P_1 and P_2 contain other overlength edges that would stop the method for clustering based on overlength edges from finding these partitions.

1) Find a plane graph from the diagram by replacing edge crossings with dummy stations.
2) Derive the dual graph, that is, the graph found from the diagram by placing a vertex in each face of the diagram. The dual graph also has an edge that cuts each edge in the diagram, and which connects the vertices in the two faces of a diagram edge. The dual graph for a diagram is unique as the diagram embedding is known.
3) Diminish the dual graph by removing unnecessary edges. These are dual graph edges that cut diagram edges of ideal edge length and dangling edges (which are dealt with using the method of Section 4.1). This leaves only the dual graph edges that cut overlength edges.
4) Partition the plane graph by finding a route through the dual graph from the vertex in the outer face, that passes through at least one other vertex and returns to the vertex in the outer face.

There are a number of possible partitions. We use a heuristic that finds a path that cuts diagram edges which are *most opposite* each other. Cuts through the diagram that consist of nearly parallel edges are more likely to result in good clusters, see Figure 15. For each new face, we take the current edge and find an overlength edge



(a) (b)

Fig. 15. Example of cuts which are likely to lead to poor and good partitions. The example in (a) is likely to lead to poor partition selection while the example in (b) is likely to lead to a better partition selection.

that is both opposite the current one in the face and as close to parallel to the starting edge as possible.

The overlength edges clustering method of Section 4.1 is still required, as the partitioning in this section only operates on edges that are in cycles formed from faces in the diagram. It is not applied to 'dangling' lines, which are common in the outer sections of the diagram, and which often have overlength edges.

5 Examples

In this section we give some example metro maps produced by our system, showing for comparison the undistorted and published maps. The time taken to generate the automatically generated maps discussed in this paper is given in Table 3. These timings were performed in Java 1.6, on a computer with a 1.4GHz Celeron M processor, 1.5GB RAM and running Windows XP. The values are the average of three runs. All maps (automatically generated, published and undistorted) can be seen in Appendix B.

TABLE 3
Time taken to generate maps.

Map	Time (In Seconds)
Atlanta	125
Bucharest	175
Mexico City	3559
Stockholm	695
Sydney	7189
Toronto	339
Washington	1237

5.1 Mexico City

The Mexico City metro map [27] is a complex, decentralized map. It has a relatively high number of lines and faces and a total of 175 stations. The officially published map is shown in Figure 16, with a version drawn in the style of the diagram layout software shown in Figure 17. The undistorted map is shown in Figure 18.

The map produced using our method is shown in Figure 19. We used the criteria weightings shown in Table 4 to produce this map. We believe our finished map shows a significant enhancement over both the undistorted and official maps. The official map has irregular station spacing and no attempt to achieve octilinear angles which is a feature of the automatically-drawn map. The labelling in the automatically-drawn map is also of good quality, particularly along long lines. The grey line to the top-right of the map shows a meander where the line has been compressed horizontally in order to fit within the bounds of the drawing area. Due to the large number of faces in this map, the clustering by partitioning algorithm was very effective in straightening a number of lines and compressing some overlength lines. A few examples of local minima are notable in our map, particularly where several lines pass through a station (where the blue, green and brown lines meet), where a triangular face exists (top middle of the diagram), or where the red and orange lines are drawn very close together (top right of the diagram). The line straightness criterion tends to force these lines to become horizontal thereby reducing the angle between them. This could be avoided by increasing the weighting for the angular resolution criterion, but in practice this tends to result in less optimal conditions elsewhere in the map.

TABLE 4
Weightings and parameters for the Mexico City and Sydney maps.

Station movement		Labelling		Other Parameters	
w_{N1}	30000.0	w_{L1}	300.0	Iterations	5
w_{N2}	50.0	w_{L2}	80.0	Pref. grid spacing l	4
w_{N3}	45.0	w_{L3}	19.0	Grid Spacing g	40
w_{N4}	220.0	w_{L4}	15.0	Min. Cluster Distance	3
w_{N5}	9250.0	w_{L5}	3.0	Max. Station Movement	8
		w_{L6}	1.0		
		w_{L7}	30.0		

5.2 Sydney

The Sydney CityRail [28] is a very large network covering an area of approximately 3600km^2 of metropolitan Sydney. The use of enlarged scale is very prominent in the central Sydney area where most of the lines converge in a tight loop around the city centre. Long horizontal lines have forced the use of diagonal labels, but all diagonal labels are of the same orientation. The official Sydney CityRail map is shown in Figure 20 and the undistorted map is shown in Figure 21. We have constrained our area of interest to the main metropolitan area of Sydney. The version of the Sydney CityRail map drawn using our method is shown in Figure 22 and uses the criteria weightings given in Table 4.

The finished version of our map has succeeded in evening out station spacing and nearly all the edges are drawn octilinearly. Labelling is also of a good quality.

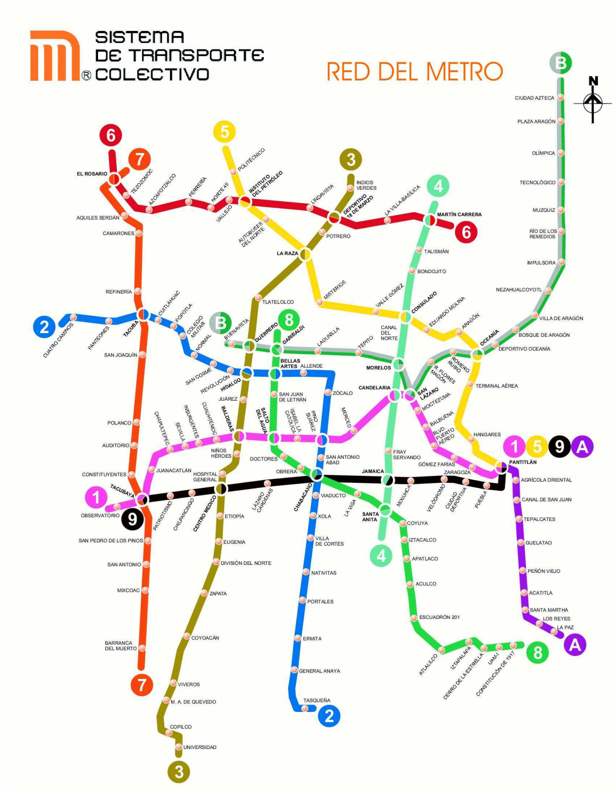


Fig. 16. Mexico City map: official layout.

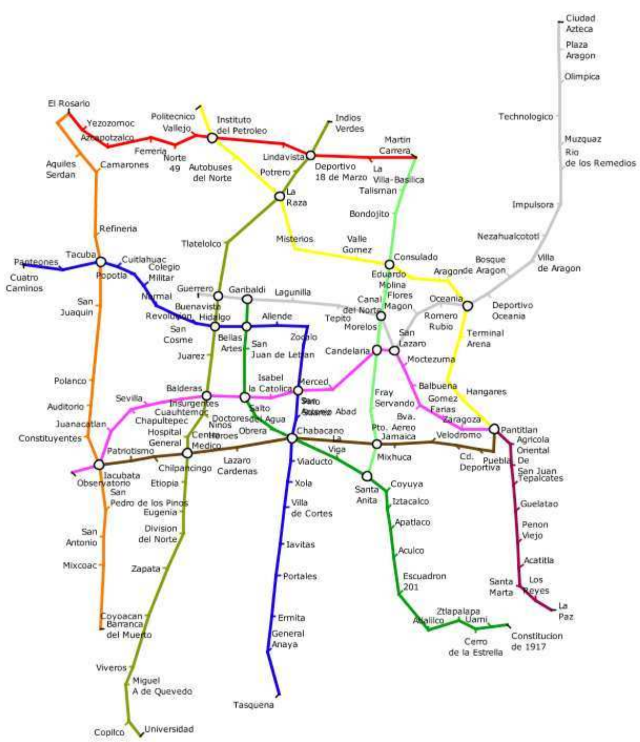


Fig. 17. Mexico City map: official layout, normalized to the layout software style.

One particular area posing a problem for our method is the central area at the right-hand side of the map. This section has up to seven lines passing through each station and features a very tight loop. The published map handles this area by significantly increasing the scale (possibly one of the most dramatic uses of enlarged scale seen in published metro maps), but our method does not explicitly handle scale enlargement for such a small area of the map. A few edges are not drawn octilinearly, most notably the bottom-most horizontal line in the map. In this case, a local minimum has been reached where none of the clustering algorithms will find the right cluster of stations as the length of some of the edges is greater than the minimum cluster distance.

6 Empirical Study

In this section, we report an empirical study conducted to evaluate maps drawn using the method layout described in Sections 3 and 4. We compared them with the official published map and undistorted map. We aimed to evaluate the following four hypotheses:

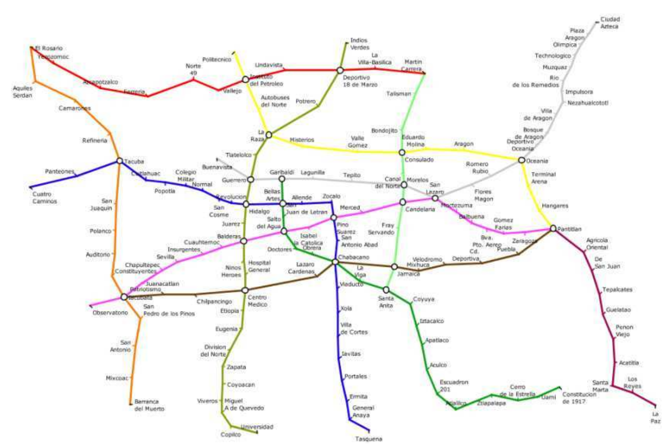


Fig. 18. Mexico City map: undistorted layout.

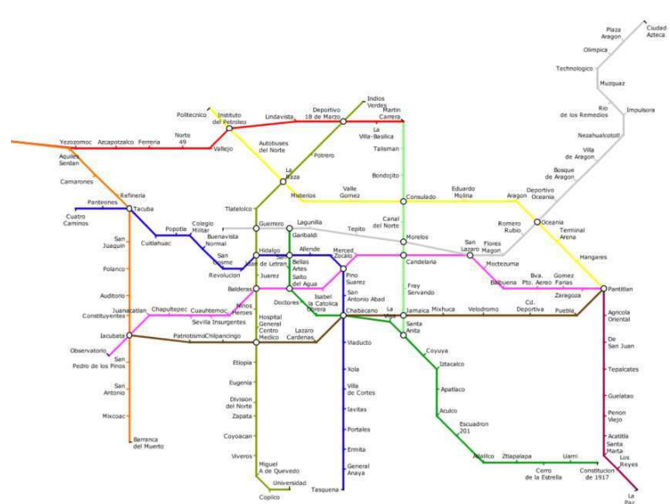


Fig. 19. Mexico City map: our layout.

A) A map of a metro system drawn with our automated software is better for finding an optimal route than a undistorted map of the system.

B) A map of a metro system drawn with our automated software is better for finding an optimal route than the official published map of the system.

C) A map of a metro system drawn with our automated software is preferred over a undistorted map of the system.

D) A map of a metro system drawn with our automated software is preferred over the official published map of the system.

The empirical experiment involves a sample of human subjects performing route-planning tasks using different map versions of the metro systems given in Table 5.

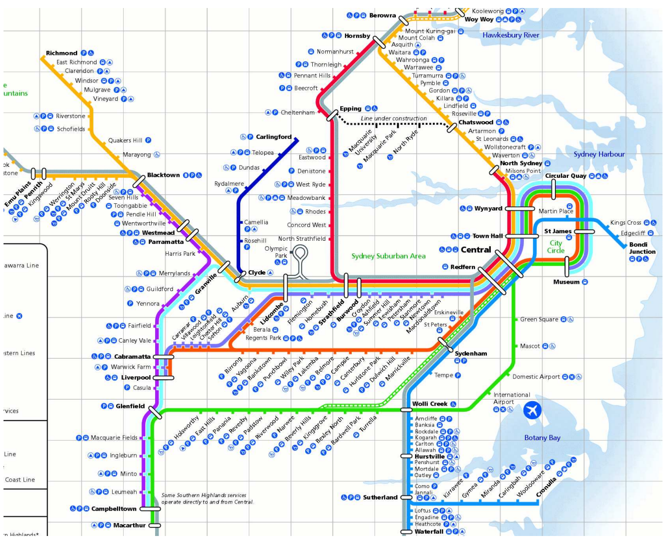


Fig. 20. Sydney CityRail map: official layout.

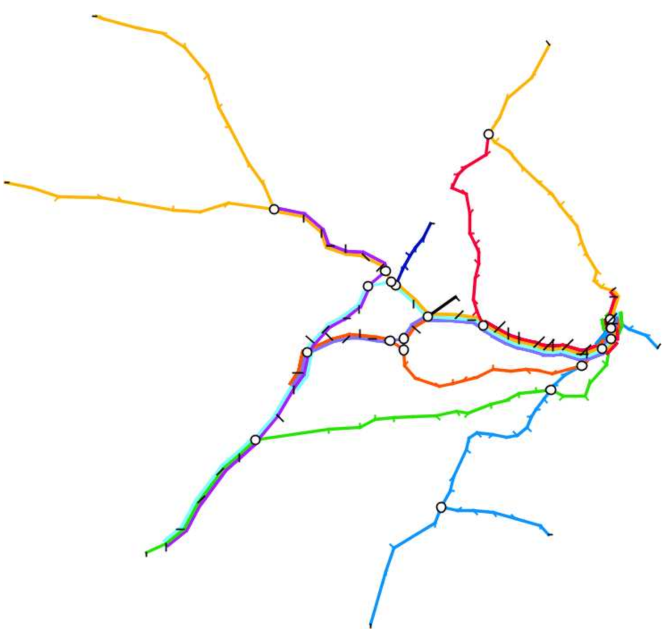


Fig. 21. Sydney CityRail map: undistorted layout.

These maps differ in characteristics and complexity from fairly simple two-line, centralised network, in the case of Atlanta, through to complex, highly interconnected, decentralised network, in the case of Mexico City.

All the maps were rendered in a similar way, to avoid discrepancies due to different label fonts or line thickness. For example, the difference between the original published Mexico City map and the normalized published map used in the empirical study can be seen in Figures 16 and 17.

TABLE 5
Metro maps used in the empirical study.

Map	Stations	Lines	Interchange	Edges Stations	Faces
1. Atlanta	39	2	3	36	1
2. Bucharest	45	3	6	45	3
3. Mexico City	175	11	24	165	19
4. Stockholm	100	3	9	101	2
5. Toronto	70	4	5	70	2
6. Washington	86	5	9	108	5

The study comprises the recording and analysis of objective and subjective measurements. Specifically, we evaluated Hypotheses A and B using objective data corresponding to the time taken by a subject to complete the task of finding a correct specific route. We evaluated Hypothesis C and D using subjective opinions obtained from the same subjects about their preferences for the maps.

6.1 Experiment design

A total of 43 subjects participated in our study, nearly all of whom were Computer Science undergraduates from the University of Kent. We split the set of subjects into three balanced groups: I, II and III.

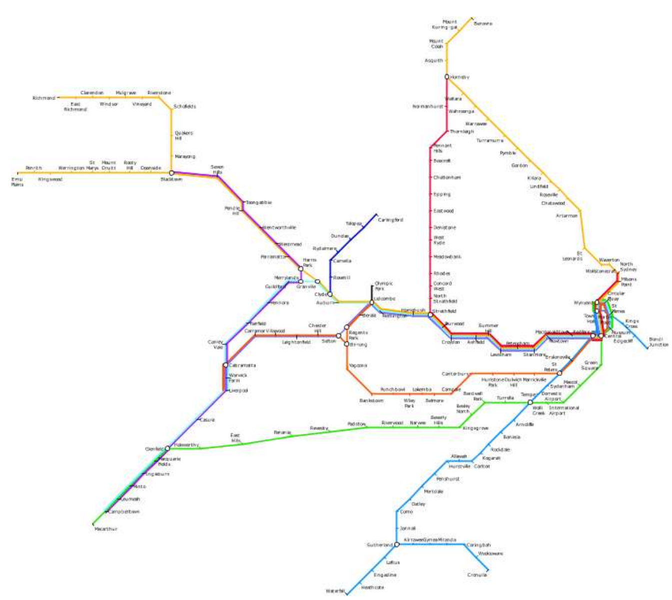


Fig. 22. Sydney CityRail map: our layout.

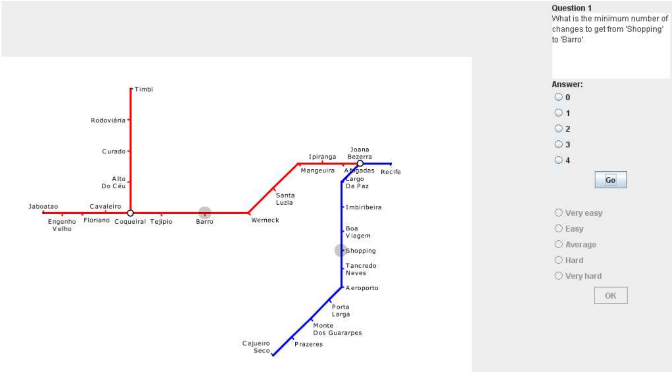


Fig. 23. Screenshot of the software application used to conduct the empirical evaluation.

Each group received exactly the same questions in exactly the same order but they only saw one variant of each map per group for each question. For example, for a question using the Atlanta map, all subjects in all groups are performing the same task but group I used the normalized published map, group II used the undistorted map and group III used the automatically-drawn map. The map variants were distributed evenly amongst the groups.

A software application was written which ensured a controlled environment when showing the maps. A screenshot of the software application is shown in Figure 23. The following procedure was used for each experiment:

i) An introductory script was read aloud and the test supervisor worked through two example questions.
ii) The subjects were told to begin and were presented with problems for 20 minutes. For each problem there was a map, a question and a list of five possible answers. The subject selected their answer and then continued on to the next problem. The subjects were able to rest after the completion of each problem.
iii) After 20 minutes had elapsed, the subjects were shown them how many of their answers were correct. For incorrect answers they were shown the right answer.
iv) The questionnaire script was then read aloud. The subjects were then shown the three variants of each map and were asked to write down their preference from "most preferable" to "least preferable".
v) The subjects were then rewarded with £5 for their time and were allowed to leave.

The full list of questions and answers for the experiments is given in Appendix B. Prior to the real experiment sessions, a pilot study was used to determine any problems in our methodology. During this pilot we were able both to find how much time would be appropriate for the number of questions we were asking and to uncover any ambiguous or impossible questions. The scripts were also refined as a result. The results from the pilot were discarded.

6.2 Statistical analysis

6.2.1 *Duration data*

Each individual in the study contributes towards learning about the duration time taken to accurately find a specific route (step (iii) of the experiment). With the data, the aim is to evaluate the time-effectiveness of each metro map using the time elapsed in completing route planning tasks. During the study we recorded "exact" measurements corresponding to accurate routes. However, in some cases, the task was performed incorrectly. However, it is natural to assume those tasks will be correctly answered with more time. The key here is that we are not assessing correctness as an outcome; rather the time taken to find the correct route. So, if an incorrect route is given at a particular time, the time taken to provide a correct route would be greater than this time. Therefore, we can and do use this as information, being compatible with the aims of the experiment without losing any information.

We used statistical methods developed in time-to-event data analysis. The idea is that we model the times when a route is given and within the model we include key parameters which allow us to assess the hypotheses A and B highlighted previously. See Appendix C for a detailed description of the model and more general statistical results.

Time-effectiveness assessment

Hypotheses A and B can be assessed (independently and jointly) using probability statements by using Bayesian methods. Effectively this allows us to evaluate Probability(A) which obviously records the estimate of the probability of hypothesis A being true. Table 6 shows the estimates of these probabilities. As we can see, the three hypotheses are highly likely since the estimated probabilities are considerably greater than 0.5.

TABLE 6
Estimated probabilities for hypotheses A and B.

Hypothesis	Probability	(Std. Err.)
A	0.986	(0.113)
B	0.994	(0.076)
A & B	0.994	(0.105)

6.2.2 *Preference data*

Table 7 displays the contingency table of individuals' preferences for map variants across metro systems (part (iv) of the experiment). The preference for the automatically-drawn map is clear, with the exception of Stockholm. Note that there are six possible rankings an individual can assign within the three map versions, which can be enumerated as follows, starting with the most preferable: (a) Auto-Un-Pub, (b) Auto-Pub-Un, (c) Un-Pub-Auto, (d) Un-Auto-Pub, (e) Pub-Auto-Un, (f) Pub-Un-Auto. Hypotheses C and D can be summarized by events (a) and (b). Therefore, assessing hypotheses C and D simultaneously (C & D) would be equivalent to evaluating how likely events (a) or (b) are. Assessing hypotheses C & D is accomplished by collapsing the six possible rankings into a dichotomous variable, defined to be 1 if the automatically-drawn metro map is the most preferable version and 0 otherwise.

TABLE 7
Contingency table of map preferences for the three map versions.

System	Preferences								
	Most			Medium			Least		
	Auto	Un	Pub	Auto	Un	Pub	Auto	Un	Pub
Atlanta	21	10	12	12	10	21	10	23	10
Bucharest	29	9	5	6	22	15	8	12	23
Mexico City	41	2	0	1	35	7	1	6	36
Stockholm	5	2	36	28	9	6	10	32	1
Toronto	24	5	14	17	20	6	2	18	23
Washington	33	0	10	9	3	31	1	40	2

We use statistical methods to model the probability of selecting the automatically-drawn version as the most preferable. See also Appendix C for details and statistical results.

Preference assessment

Hypotheses C and D can be statistically assessed by comparing estimated probabilities. Estimated probabilities are displayed in Table 8. Here we can see that the automatic map is the more preferable to the alternatives in Mexico City and Washington. However, we see that for Stockholm the published map remains preferable. Overall, the automatically-drawn metro map is the most preferable, since the overall estimated probability to choose this version as the most preferable is equal to 0.597. See Appendix C for further details.

TABLE 8
Estimated probabilities for hypotheses C and D by metro map system.

System	Probability	(Std. Err.)
Atlanta	0.439	(0.496)
Bucharest	0.990	(0.011)
Mexico City	0.993	(0.007)
Stockholm	0.008	(0.003)
Toronto	0.778	(0.414)
Washington	0.995	(0.009)

6.3 Study Summary

In this study, we found some quantitative evidence to assert that the automatically-drawn metro map version helps users to find routes more efficiently than the alternatives. Moreover, we found that preferences tend to favour the automatically-drawn map version, particularly in highly complex metro systems, such as Mexico City and Washington. However, the evidence is not conclusive due to the Stockholm metro system. Further analysis would be required to establish individuals' preferences in terms of metro map complexity. However, the automatically-drawn metro maps have been shown to be the best maps in terms of effectiveness of finding correct routes.

7 Conclusions

We have described an automatic metro map layout system based on multicriteria optimization. The method includes labelling and station clustering. Our empirical study shows that, for some of the maps tested, the layouts produced by the method can be considered better for route planning than both published and undistorted layouts.

Whilst improving on some less effective published maps, it is unlikely that this system will generate maps that are better than the best hand drawn maps. The complexity of the maps, combined with the sophisticated decisions made by map designers, mean that many published maps have features that are not adequately covered by our optimizer. These include: avoiding label overlap, multiple parallel lines and local enlargement. Whilst aesthetic criteria could be developed to deal with these circumstances, there are always likely to be further issues of a more subtle nature.

In terms of future work, various aspects of the layout research could be taken forward, as the current layout mechanism does not fully capture all aspects of published metro maps. Line bends in between stations could be added, as currently changes in line direction are only allowed at stations. In addition, other geographic features such as rivers, shoreline and parkland could be shown on the diagram. Aesthetic criteria could be added to integrate these features in the layout.

The optimizer has great potential for improvement. The criteria weighting is performed in an ad-hoc manner, and whilst a more systematic method for deciding weighting is difficult to design, further empirical study of diagrams drawn with different characteristics would provide evidence for improving weightings. Also, if particular weightings lead to an automatic drawing similar to a published map (perhaps adapting a layout by example approach [29]), it might be possible to characterize the map in terms of weighting, enabling other diagrams to be drawn in the style of that map.

In addition, the performance of the optimizer is slow, and little effort has been put into improving the computation time. Large speed ups are possible by: integrating the calculation of the metrics (which often perform very similar item-item comparisons, and so repeated iterations might be avoided); avoiding the comparison of items that are far away from each other in the diagram;

and reuse of calculations from previous iterations where items have not moved. Some criteria might also be removed, for example if a the edge length criterion were measured as the square of difference between the current edge length and desired edge length, then this might remove the need for the balanced edge length criterion.

Finally, applying this work beyond the layout of complete maps is feasible. The frequent use of the metro map metaphor in laying out non-transport based information means that there is opportunity to provide an automated layout mechanism for such areas. In addition, the wide spread use of small devices connected to the internet, such as mobile phones, means that the provision of personal travel maps would seem to be a promising application for automatic metro map layout.

ACKNOWLEDGMENTS

The first author was supported by an ESPRC PhD studentship. The third author was supported by CONACyT (Mexico), scholarship 159977/229515.

REFERENCES

[1] "Transport for London website," http://www.tfl.gov.uk/tube/maps/, 22 Apr. 2005.
[2] K. Garland, *Mr. Beck's Underground Map*. Capital Transport Publishing, 1994.
[3] S. Anand, S. Avelar, J. Ware, and J. Jackson, "Automated Schematic Map Production using Simulated Annealing and Gradient Descent Approaches," in *15th Annual GIS Research UK Conference (GISRUK 2007) Dublin*, 2007.
[4] R. A. Burkhard and M. Meier, "Tube map: evaluation of a visual metaphor for interfunctional communication of complex projects," in *4th Int. Conference on Knowledge Management (I-KNOW 2004)*. Graz, Austria: Know-Centre Graz, 2004, pp. 449–456.
[5] W. C. Hahn and R. A. Weinberg, "A subway map of cancer pathways," *Nature Reviews Cancer*, vol. 2, no. 5, May 2002.
[6] T. Bang, K. Grønbæk, and P. S. Hansen, "Using a metro map metaphor for organizing web-based learning resources," in *ED-MEDIA 2002*. AACE, 2002, pp. 647–652.
[7] J. M. Stott and P. Rodgers, "Metro map layout using multicriteria optimization," in *Informationan Visualization (IV04)*. London, England: IEEE, 2004, pp. 355–362.
[8] ——, "Automatic metro map design techniques," in *22nd Int. Cartographic Conference*, A Coruña, 2005.
[9] J. M. Stott, "Automatic layout of metro maps using multicriteria optimisation," Ph.D. dissertation, Computing Laboratory, University of Kent, 2008.
[10] S.-H. Hong, D. Merrick, and H. A. D. do Noscimento, "The metro map layout problem," in *APVIS 2004*, 2004, pp. 91–100.
[11] ——, "Automatic visualisation of metro maps," *J. Visual Languages and Computing*, vol. 17, pp. 203–224, 2006.
[12] M. Nöllenburg and A. Wolff, "A Mixed-Integer Program for Drawing High-Quality Metro Maps," in *13th Int. Symp. on Graph Drawing (GD2005)*, ser. LNCS, vol. 3843, 2005, pp. 321–333.
[13] D. Merrick and J. Gudmundsson, "Path simplification for metro map layout," in *14th Int. Symp. on Graph Drawing (GD2006)*, ser. LNCS, vol. 4372. Springer-Verlag, 2007, pp. 258–269.
[14] M. A. Bekos, M. Kaufmann, K. Potika, and A. Symvonis, "Line Crossing Minimization on Metro Maps," in *15th Int. Symp. on Graph Drawing (GD2007)*, ser. LNCS, vol. 4875, 2007, pp. 231–242.
[15] M. Agrawala and C. Stolte, "Rendering effective route maps: improving usability through generalization," in *SIGGRAPH 2001*, Los Angeles, USA, 2001, pp. 241–249.
[16] S. Avelar and R. Huber, "Modeling a public transport network for generation of schematic maps and location queries," in *20th Int. Cartographic Conference*, 2001, pp. 1472–1480.
[17] H. Casakin, T. Barkowsky, A. Klippel, and C. Freska, "Schematic maps as wayfinding aids," in *Spatial Cognition II*, ser. LNCS, no. 1849. Springer-Verlag, 2001, pp. 54–71.
[18] P. Yates and M. Humphries, "The generation of schematic diagrams from geographic representations of networks," in *The Spatial Information Research Centre's 10th Colloquium*, 1998.
[19] S. Cabello, M. de Berg, S. van Dijk, and M. J. van Kreveld, "Schematization of road networks," in *17th Symp. on Computational Geometry*. ACM, 2001, pp. 33–39.
[20] U. Lauther and A. Stübinger, "Generating schematic cable plans using spring embedder methods," in *Graph Drawing 2001*, ser. LNCS, vol. 2265. Springer-Verlag, 2001, pp. 465–466.
[21] R. Davidson and D. Harel, "Drawing graphs nicely using simulated annealing," *ACM Transactions on Graphics*, vol. 15, no. 4, pp. 301–331, 1996.
[22] C. Kosak, J. Marks, and S. Shieber, "A parallel genetic algorithm for network-diagram layout," in *4th Int. Conference on Genetic Algorithms*, 1991, pp. 458–465.
[23] H. C. Purchase, R. F. Cohen, and M. Jones, "Validating graph drawing aesthetics," in *Symp. on Graph Drawing*, ser. Lecture Notes in Computer Science (LNCS), F. J. Brandenburg, Ed., vol. 1027. Passau, Germany: Springer-Verlag, 1995, pp. 435–446.
[24] T. Barkowsky, L. J. Latecki, and K.-F. Richter, "Schematizing maps: simplification of geographic shapes by discrete curve elimination," in *Spatial Cognition II*, ser. LNCS, vol. 1849. Springer, 2000, pp. 41–53.
[25] M. Nöllenburg, "Automated drawing of metro maps," Master's thesis, Institut für Theoretische Informatik, Universität Karlsruhe, 20 Aug. 2005.
[26] T. Slocum, R. McMaster, F. Kessler, and H. Howard, *Thematic Cartography and Geographic Visualization*. Prentice Hall, 2005.
[27] "Mexico City 'Sistema de Transporte Colectivo' website," http://www.metro.df.gob.mx/, 13 Feb. 2008.
[28] "Sydney CityRail website," http://www.cityrail.info/, 1 Sep. 2007.
[29] I. Vollick, D. Vogel, M. Agrawala, and A. Herzmann, "Specifying Label Layout Styles by Example," in *Proceedings of the 20th ACM Symposium on User Interface Software and Technology (UIST07)*. ACM, October 2007, pp. 221–230.

Jonathan Stott graduated from the University of Kent with a first class degree with honours in Computer Science in 2003. In 2008 he completed a PhD at the same institution with a thesis titled "Automatic Layout of Metro Maps using Multicriteria Optimisation".

Peter Rodgers is a Senior Lecturer at the University of Kent. His main research interests are in diagrammatic visualization, including graph and Euler diagram layout techniques. He has led several research projects supported by national and international funding bodies.

Juan Carlos Martínez-Ovando is a Postgraduate Student in Statistics at the University of Kent. His research focuses on Bayesian nonparametric methods.

Stephen G. Walker is a Professor in Statistics at the University of Kent. His main area of research is Bayesian nonparametric methods. He has served as a program committee member for several conferences in Bayesian statistics.

Appendices for Automatic Metro Map Layout Using Multicriteria Optimization

Appendix A
Effects of the Station Movement Criteria

In this section we demonstrate the effectiveness of the station movement criteria and clustering by showing an example layout with all criteria. Then, for each criteria or clustering method, we show the same example but with that criteria given a weighting of zero. Figure 1 shows Stockholm drawn with all the criteria given in Section 3.2 and weightings as in Table 1. Removing angular resolution as shown in Figure 2, leads to fewer 90 degree angles, and more 45 degree angles. Removing the edge length criterion as shown in Figure 3, produces a diagram with very uneven station spacing. Figure 4 demonstrates the effect of removing the balanced edge length criterion, and again, station spacing is uneven, but the problematic edges are often different to those produced when the edge length criterion is removed. Removing the line straightness criterion, as shown in Figure 5, produces a diagram with jagged edges. With the final criteria, octilinarity, removed as shown in Figure 6 the resultant diagram has fewer lines that are horizontally, vertically or diagonally aligned.

For clustering, Figure 7 shows the effect of not applying the line straightness clustering method to the Stockholm map. The diagram has extra line bends on most of the lines compared to the diagram with the clustering method included. Removing the overlength edge clustering method results in the diagram shown in Figure 8, which has numerous line sections that have considerable larger gaps between stations than when the clustering method is included. Finally, to demonstrate the effectiveness of the partitioning clustering method, we must consider another metro map, as Stockholm only has one inner face, and multiple inner faces are required for partitioning to have an effect. Instead, Bucharest with all criteria and clustering is given in Figure 9, drawn with the weightings as in Table 1. In contrast, when the partitioning clustering method is switched off as in Figure 10, the red line around the inner faces is increased in length, with more overlength edges.

TABLE 1
Criteria weightings and parameters used for the Stockholm and Bucharest maps.

Station movement		Other Parameters	
w_{N1}	30000.0	Iterations	6
w_{N2}	50.0	Pref. grid spacing l	4
w_{N3}	45.0	Grid Spacing g	30
w_{N4}	220.0	Min. Cluster Distance	3
w_{N5}	9250.0	Max. Station Movement	8

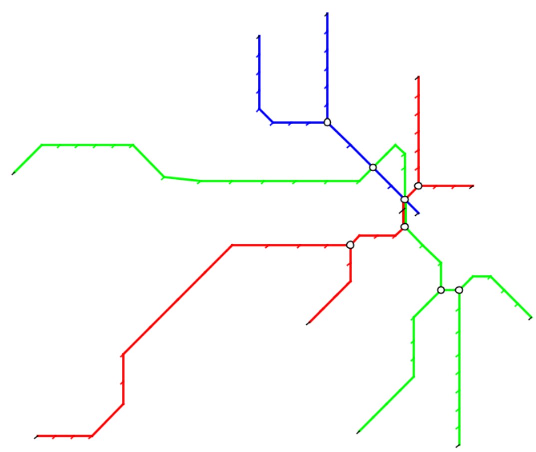


Fig. 1. Stockholm lines and stations with all criteria included.

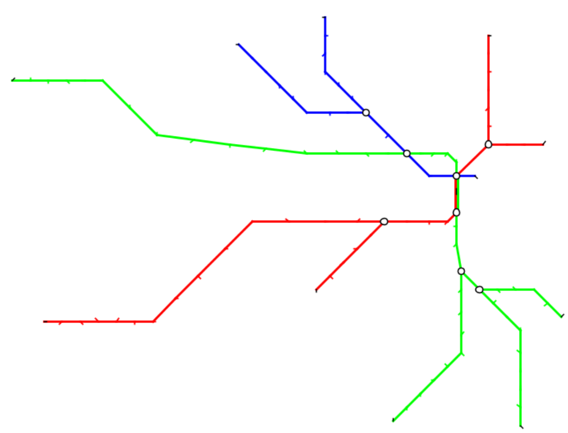


Fig. 2. Stockholm with no angular resolution.

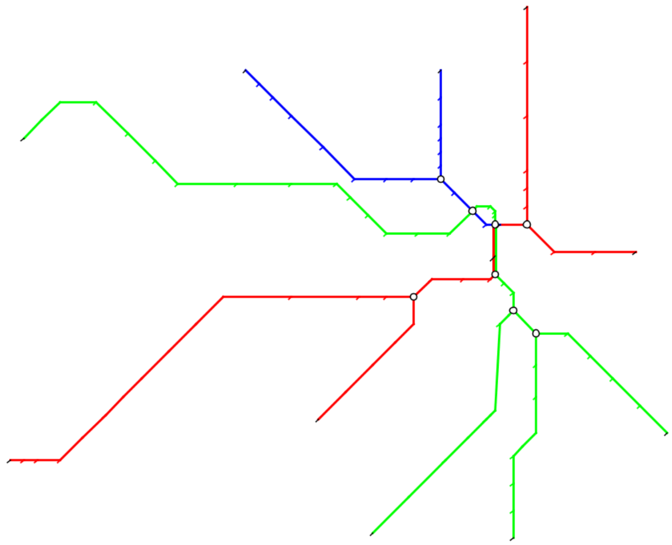


Fig. 3. Stockholm with no edge length.

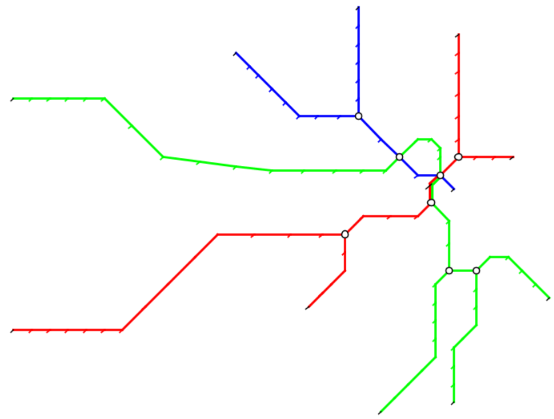


Fig. 4. Stockholm with no balanced edge length.

APPENDIX B
FULL LIST OF STUDY QUESTIONS AND MAPS

B.1 Questions

Tables 1-6 give the list of questions asked about the maps (which were drawn with the weightings in Table 1), along with the correct answers.

B.2 Atlanta

- Undistorted map (Figure 11).
- Normalized published map (Figure 12).
- Automatically-drawn map (Figure 13).

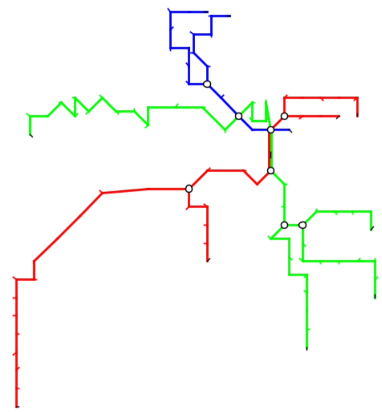


Fig. 5. Stockholm with no line straightness.

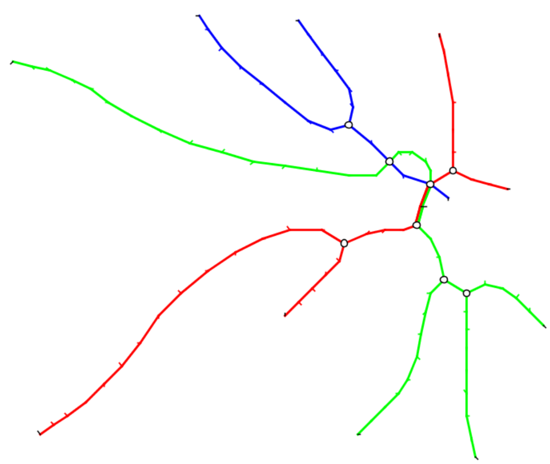


Fig. 6. Stockholm with no octilinearity.

TABLE 2
Atlanta questions.

Number	Question	Answer Options	Answer
1.a	How many stations do you pass through to get from 'Bankhead' to 'Lenox'	$\{8, 9, 10, 11, 12\}$	10
1.b	How many stations do you pass through to get from 'Vine City' to 'Midtown'	$\{4, 5, 6, 7, 8\}$	5
1.c	What is the minimum number of changes to get from 'Indian Creek' to 'West Lake'	$\{0, 1, 2, 3, 4\}$	0
1.d	What is the minimum number of changes to get from 'College Park' to 'Five Points'	$\{0, 1, 2, 3, 4\}$	0

TABLE 3
Bucharest questions.

Number	Question	Answer Options	Answer
2.a	How many stations do you pass through to get from 'Pacii' to 'Titan'	$\{9, 10, 11, 12, 13\}$	10
2.b	How many stations do you pass through to get from 'Romana' to 'Piata Sudului'	$\{1, 2, 3, 4, 5\}$	5
2.c	What is the minimum number of changes to get from 'Eroilor' to 'Iancului'	$\{0, 1, 2, 3, 4\}$	0
2.d	What is the minimum number of changes to get from 'Eroii Revolutiei' to '1 Mai'	$\{0, 1, 2, 3, 4\}$	2

TABLE 4
Mexico City questions.

Number	Question	Answer Options	Answer
3.a	How many stations do you pass through to get from 'Balderas' to 'Consulado'	$\{4, 5, 6, 7, 8\}$	7
3.b	How many stations do you pass through to get from 'Refineria' to 'Patriotismo'	$\{4, 5, 6, 7, 8\}$	6
3.c	What is the minimum number of changes to get from 'Sevilla' to 'Aragon'	$\{0, 1, 2, 3, 4\}$	2
3.d	What is the minimum number of changes to get from 'Martin Carrera' to 'La Paz'	$\{0, 1, 2, 3, 4\}$	2

TABLE 5
Stockholm questions.

Number	Question	Answer Options	Answer
4.a	How many stations do you pass through to get from 'Stora Mossen' to 'Karlaplan'	$\{6, 7, 8, 9, 10\}$	10
4.b	How many stations do you pass through to get from 'Liljeholmen' to 'Kungstrådgarden'	$\{2, 3, 4, 5, 6\}$	6
4.c	What is the minimum number of changes to get from 'Kista' to 'T-Centralen'	$\{0, 1, 2, 3, 4\}$	0
4.d	What is the minimum number of changes to get from 'Bergamossen' to 'Axelsburg'	$\{0, 1, 2, 3, 4\}$	1

TABLE 6
Toronto questions.

Number	Question	Answer Options	Answer
5.a	How many stations do you pass through to get from 'Dupont' to 'Chester'	$\{6, 7, 8, 9, 10\}$	7
5.b	How many stations do you pass through to get from 'Lansdowne' to 'York Mills'	$\{10, 11, 12, 13, 14\}$	14
5.c	What is the minimum number of changes to get from 'Bayview' to 'Union'	$\{0, 1, 2, 3, 4\}$	1
5.d	What is the minimum number of changes to get from 'Midland' to 'Dundas'	$\{0, 1, 2, 3, 4\}$	2

TABLE 7
Washington questions.

Number	Question	Answer Options	Answer
6.a	How many stations do you pass through to get from 'Pentagon' to 'Court House'	$\{1, 2, 3, 4, 5\}$	2
6.b	How many stations do you pass through to get from 'Cleveland Park' to 'Federal Triangle'	$\{4, 5, 6, 7, 8\}$	4
6.c	What is the minimum number of changes to get from 'Metro Center' to 'Takoma'	$\{0, 1, 2, 3, 4\}$	0
6.d	What is the minimum number of changes to get from 'Largo Town Center' to 'Eisenhower Ave.'	$\{0, 1, 2, 3, 4\}$	2

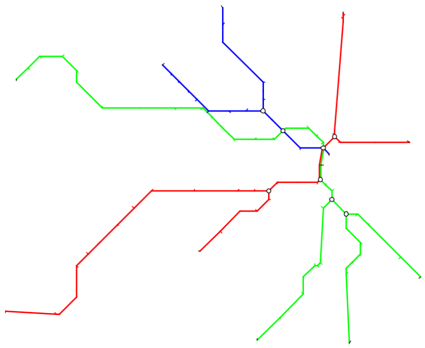


Fig. 7. Stockholm with no non-straight line clustering.

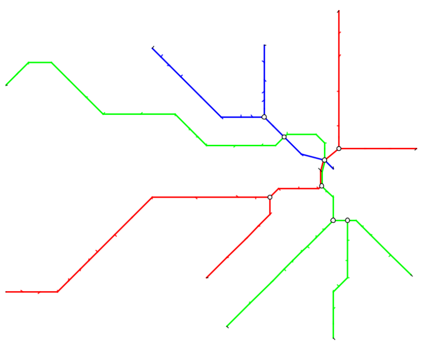


Fig. 8. Stockholm with no overlength edge clustering.

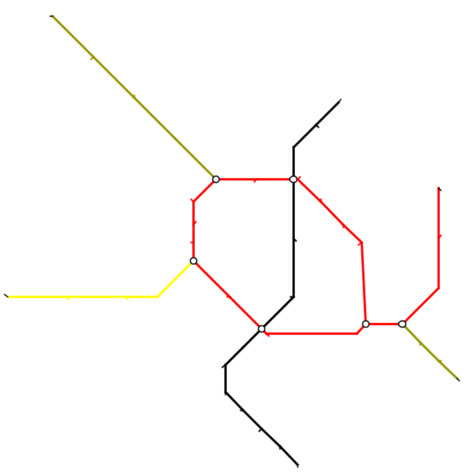


Fig. 9. Bucharest lines and stations with all criteria and clustering included.

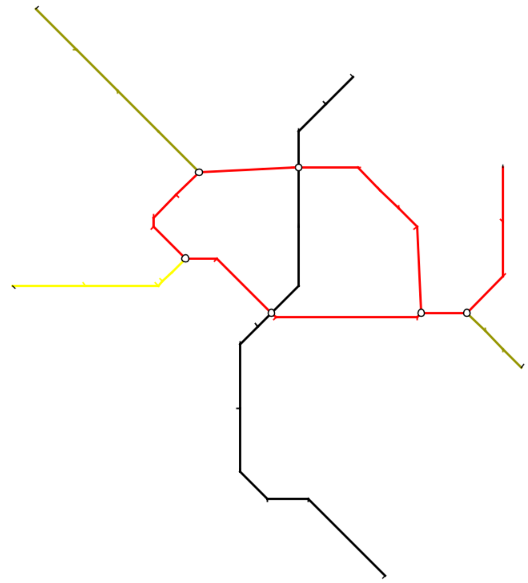


Fig. 10. Bucharest with no partitioning clustering included.

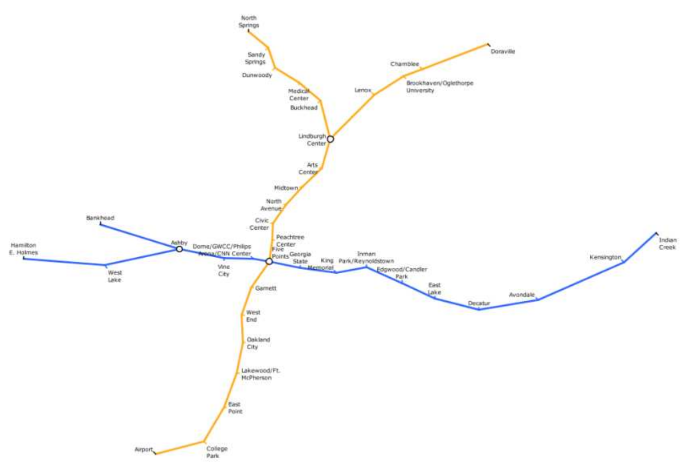


Fig. 11. Atlanta MARTA undistorted map.

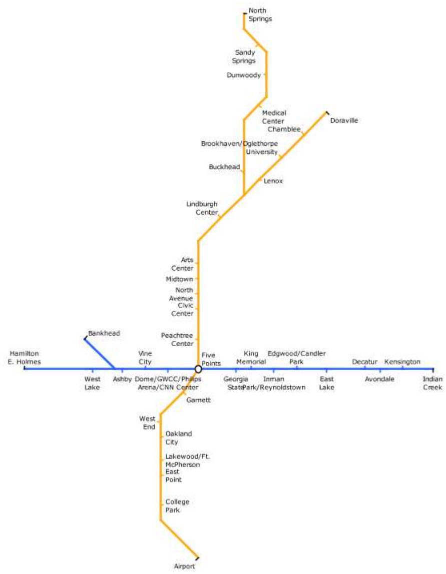


Fig. 12. Atlanta MARTA normalized published map.

Fig. 13. Atlanta MARTA automatically-drawn map.

B.3 Bucharest

- Undistorted map (Figure 14).
- Normalized published map (Figure 15).
- Automatically-drawn map (Figure 16).

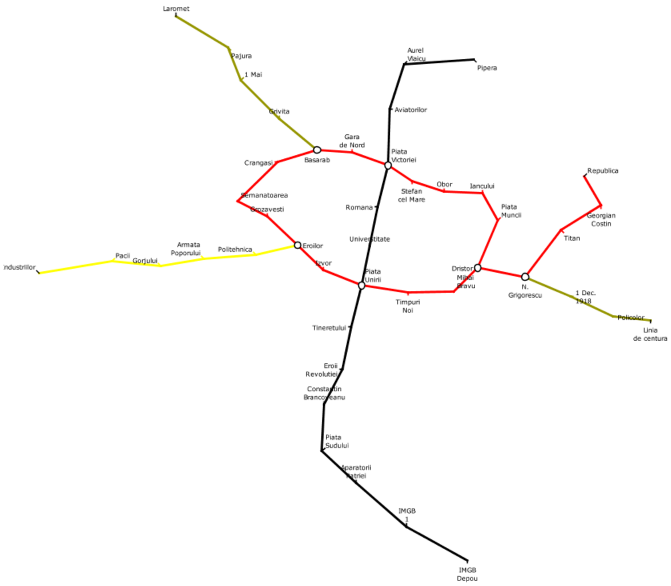


Fig. 14. Bucharest undistorted map.

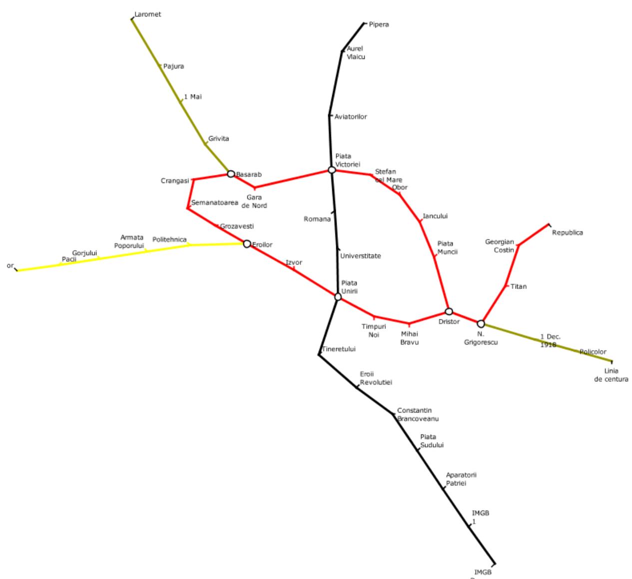


Fig. 15. Bucharest normalized published map.

Fig. 16. Bucharest automatically-drawn map.

B.4 Mexico City

- Undistorted map (Figure 17).
- Normalized published map (Figure 18).
- Automatically-drawn map (Figure 19).

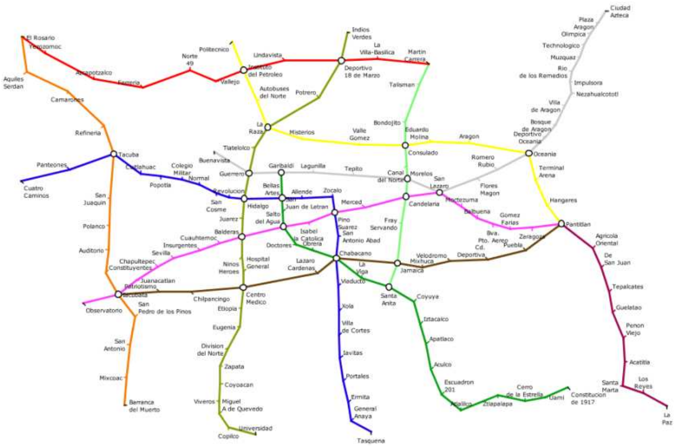


Fig. 17. Mexico City undistorted map.

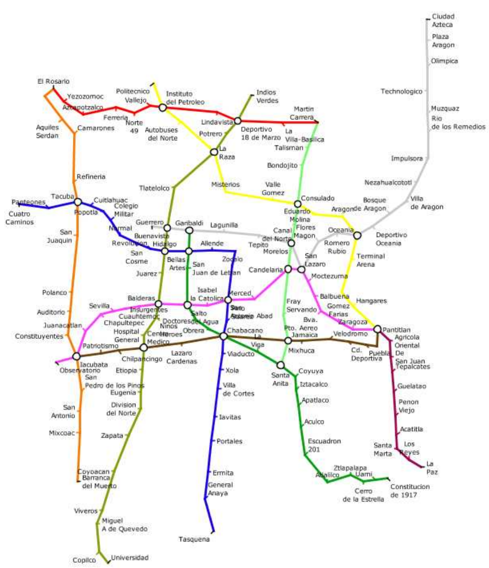


Fig. 18. Mexico City normalized published map.

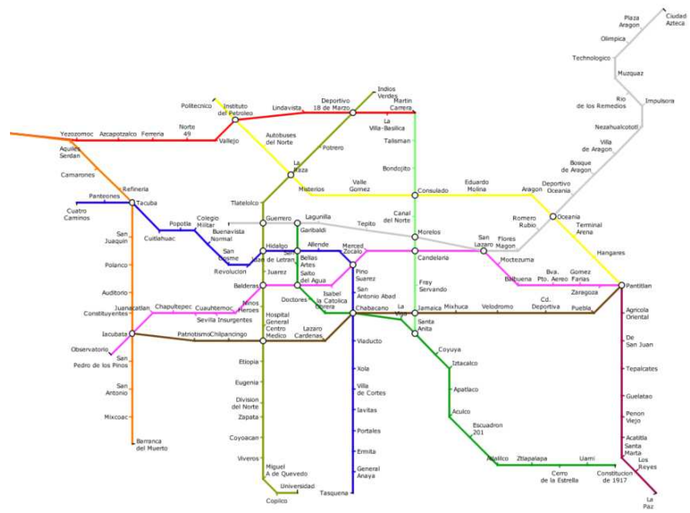


Fig. 19. Mexico City automatically-drawn map.

B.5 Stockholm

- Undistorted map (Figure 20).
- Normalized published map (Figure 21).
- Automatically-drawn map (Figure 22).

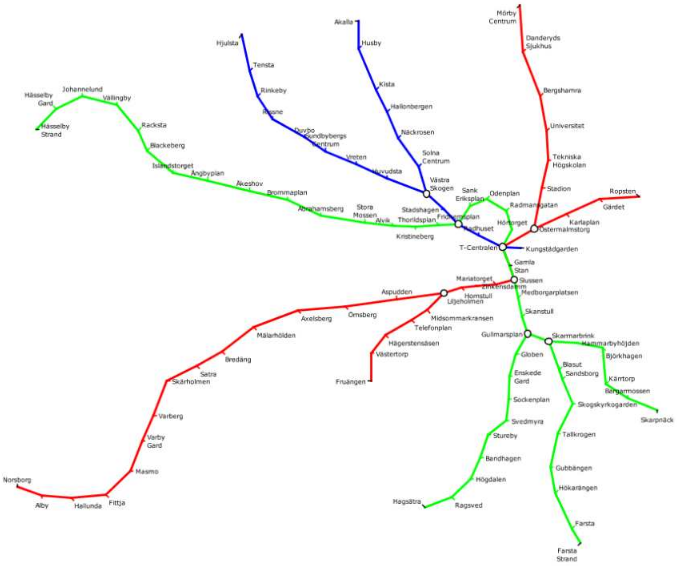


Fig. 20. Stockholm undistorted map.

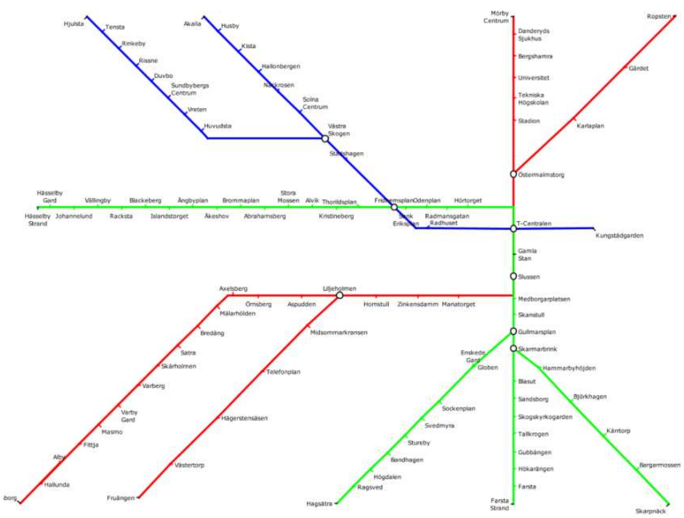


Fig. 21. Stockholm normalized published map.

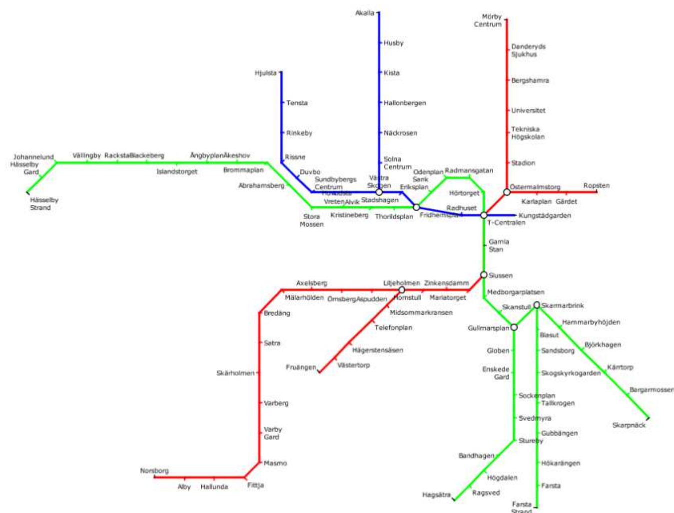


Fig. 22. Stockholm automatically-drawn map.

B.6 Toronto

- Undistorted map (Figure 23).
- Normalized published map (Figure 24).
- Automatically-drawn map (Figure 25).

Fig. 23. Toronto undistorted map.

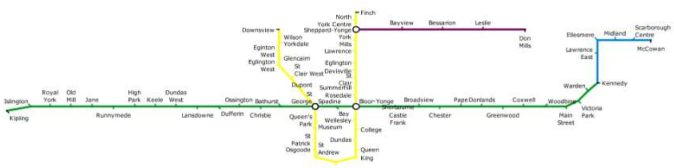


Fig. 24. Toronto normalized published map.

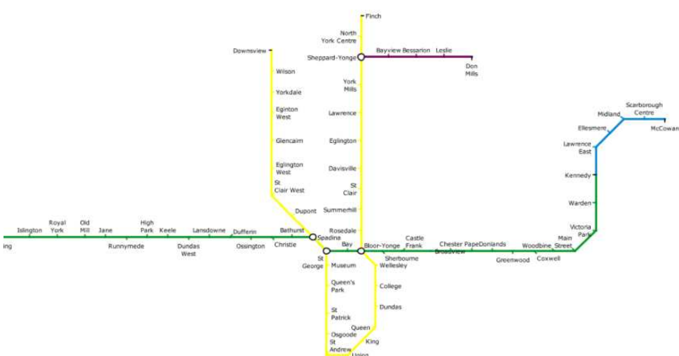


Fig. 25. Toronto automatically-drawn map.

B.7 Washington D.C.

- Undistorted map (Figure 26).
- Normalized published map (Figure 27).
- Automatically-drawn map (Figure 28).

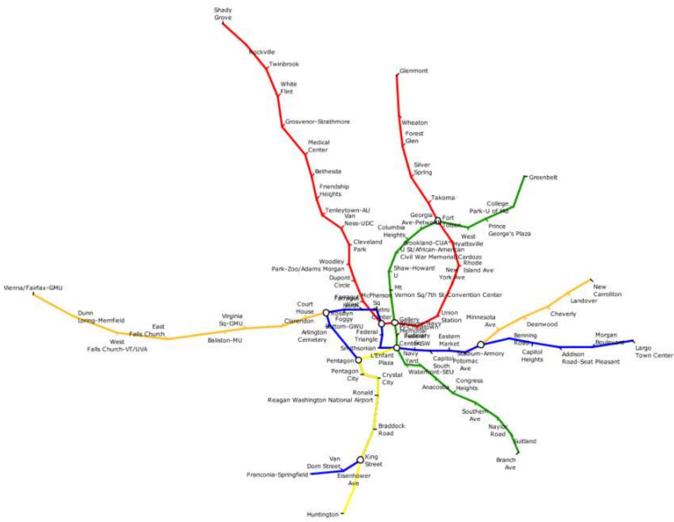


Fig. 26. Washington D.C. undistorted map.

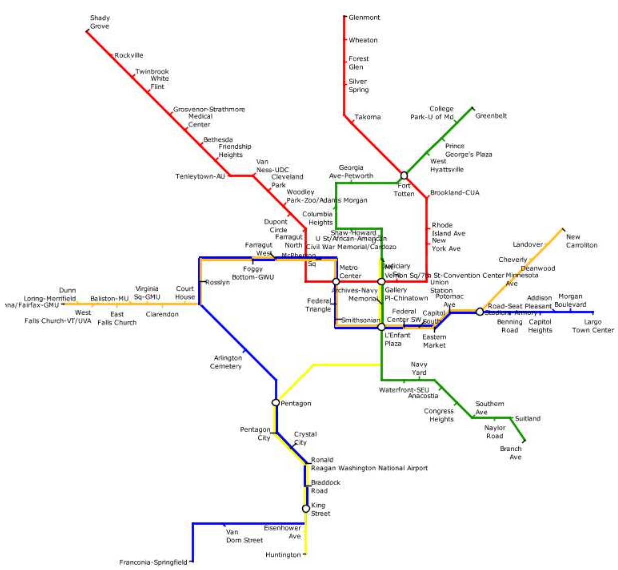


Fig. 27. Washington D.C. normalized published map.

B.8 Associated Material

As a number of sessions of the empirical experiment were run at different times, it was essential that each session was as similar as possible. The following scripts and handouts were standard across all sessions.

B.9 Preliminary Script

This script was read aloud to candidates before they started the experiment.

[The test software should be shown on the projector]

Please do not start using a computer until told to do so. During this test, do not talk, or attempt to see what other participants are answering. If you have a query, please raise your hand.

Although we ask for your login so that we can collate the data, the results of this test and questionnaire will be anonymized.

You will be presented with a sequence of metro map diagrams. The test will pose a question for each map which requires an answer to be selected. The question will require you to plan a route between two stations on the map.

You will first need to enter your login, level of study, year of study, age and gender and click OK. Do not do this yet, you will be told when to start the test.

[Enter login test level of study Undergraduate, year of study 2, Age 25, Gender Male, then press OK. Press Start]

A metro map is used to depict the interconnections on a public transport system so that the user is able to plan and undertake a specific journey. Stations are represented by circles which are labelled with the name of the station. A line in a single colour indicates which stations are connected by direct services. Where two or more lines pass through a single station you are able to change from one line to the other. See an example of a metro map on the projector.

When you start the test, you will be shown a metro map in the main part of the screen [Point to map]. The question will be shown in the top-right corner [Point to question] with a selection of answers below [Point to answers].

The questions that will be asked will involve

Fig. 28. Washington D.C. automatically-drawn map.

planning a journey between two stations. The stations will be highlighted on the map in order that you can identify them more easily [Point to highlighted stations].

Once you have worked out the route for the question, click the button next to the answer in the list shown before clicking the Go button.

In this case the question is How many changes are required to get from 'Shopping' to 'Barro'. I can see that you need to change once from one coloured line to another [point on screen]. So I click 1 [Click option 1]. Then I click GO [Click GO].

After each question you need to give an indication of the difficulty of the question. To do this, select the appropriate option from the list from very easy to very hard. Then click the OK button.

The period between clicking Go and answering the difficulty question are an opportunity to rest, if you need to do so, as timing does not start again until the OK button is clicked.

[Click Average then OK]

This next question asks me How many stations do I go through to get from Aeroporto to Santa Luzia. With this sort of question, you do not count the end stations, only the stations in-between. Counting the stations, including the station that requires me to change I get an answer of 9 [Point at each intermediate station, counting]. [Click option 9, then Go]

Please do not rush the questions, and take some effort to get the questions correct. Whilst we are measuring the time it takes to complete each answer, we do not mind if you do not complete all the questions.

You will be presented with questions for 20 minutes. After this time is up you will be shown how many questions you got correct as well the answers to any questions that you got incorrect.

At the end of the test, do not log off.

After the test, you need to complete a short questionnaire

Enter your details and press OK then the start button to begin the test now.

B.10 Postliminary Script

This script was read out after the interactive part of the experiment had concluded.

[The first slide should be showing on the projector]

[Hand out 1 questionnaire and 1 pen to each participant]

Please first fill in your login on the sheet in front of you, and then look up at the projector screen. You will be shown three metro maps at a time. Please decide which of these maps would be best for navigating a metro map system. As each slide is shown, write down 1, 2 or 3 in the spaces below, where 1 is the most preferable map and 3 is the least preferable map

I will count down from 5 before showing the next set of metro maps

[Count down 5, 4, 3, 2, 1, then show the next slide]

[Wait for a minute and then count down 5, 4, 3, 2, 1, for each slide, until the end of the presentation appears]

Please now take 5 minutes to fill in questions 1, 2, 3 and 4 by hand.

[Wait for 5 minutes and then start handing out five pound notes, getting signatures and handing out debriefing scripts. The experiment is now over, and you can answer questions about the tasks].

B.11 Questionnaire

This is the questionnaire that each candidate was asked to fill in in relation to the postliminary script (Section B.10).

Your Login:

Please first fill in your login above, and then look up at the projector screen. You will be shown three metro maps at a time. Please decide which of these maps would be best for navigating a metro map system. First write down your login, then, as each slide is shown, write down 1, 2 or 3 in the spaces below, where 1 is the most preferable map and 3 is the least preferable map.

TABLE 8
Preference grid.

Slide	Preference for Map A	Preference for Map B	Preference for Map C
Slide 1			
Slide 2			
Slide 3			
Slide 4			
Slide 5			
Slide 6			

Please now take five minutes to answer the questions below

1) Have you seen any of the metro maps shown here before these tests? If so, which ones?
2) Which features of the metro map layout did you find most helpful when completing the tests?
3) Which features of the metro map layout did you find least helpful when completing the tests?
4) Did you find any of the questions ambiguous?

B.12 Concluding Handout

This text was given to candidates after they had completed the questionnaire and before they left the experiment.

> Thank you for participating in this research
>
> You were presented with a number of maps which were drawn using three different techniques. One version was drawn using the geographic layout of stations; the second version was drawn from a published map of the network; the final version was drawn using an automatic method which balances aesthetic criteria to try and find an optimal layout.
>
> The purpose of this research is to qualify some design aesthetics for automatically laying out metro maps and to see if our automated method is good at producing comprehensible diagrams. The idea is that being able to automatically produce metro maps might improve their use for navigating metro networks. In addition, being able to automatically generate the such maps could lead to them being more widely used for many other application areas.
>
> We would appreciate it if you did not discuss this experiment with other students in the university. These experiments will be continuing through the last two weeks of term, and having subjects who have prior knowledge of what the tests are about makes the data less useful.
>
> Thank you again for your contribution.

APPENDIX C DETAILED STATISTICAL ANALYSIS OF THE EMPIRICAL STUDY

In this section, we give more detail of the statistical analysis summarized in Section 6.2.

C.1 Duration data

Each subject in the study contributes towards learning about the duration time taken to correctly find a specific route (corresponding to step (iii) in the experiment). We used statistical methods developed in time–to–event data analysis, such as survival analysis, to study the duration times. More specifically, we used a proportional hazards [1] model with frailties [2]. The idea here is that T, the duration time, is modeled differently for each individual and each map, so there are 2 indices for each T_{ij}, i indexing the individual and j indexing the map. If for the moment we use ψ to denote the parameters of interest to be estimated then we will be basing estimation on the likelihood function

$$L = \prod_{i,j} f(t_{ij}|\psi)$$

where t_{ij} are the observed T_{ij}, and f is the density function of the duration times. Values of ψ which make this likelihood large are considered good values. The one which maximises the likelihood is often taken as an estimate of ψ. So we need to specify a $f(t|\psi)$. We develop this model using a proportional hazard model with frailties. This involves working with the hazard function, which depends on f via

$$h(t) = f(t)/S(t),$$

where $S(t)$ is the survival function given by $S(t) = \int_t^\infty f(t)dt$. Hence, $f(t) = h(t)S(t)$. The hazard function for observation t_{ij} is given by

$$h\left(t_{ij}\right) = h_0\left(t_{ij}\right)\exp\left(\theta_i + \boldsymbol{x}'_{ij}\boldsymbol{\beta}\right), \tag{1}$$

where h_0 stands for the underlying (baseline) hazard function which would be the hazard function for all observations if every individual was "the same" and the maps were all the same. Here, $\theta_i = \log \upsilon_i$ denotes the multiplicative random effect for repeated observations of the ith subject; that is, some effect on t_{ij} provided by individual i. The υ_is, known as frailties, are assumed to be independent and identically distributed random variables. The fixed effects associated with the map and map type are given by

$$\boldsymbol{x}'_{ij}\boldsymbol{\beta} = \beta_{\text{auto}}\mathbf{1}_{\text{auto}}(ij) + \beta_{\text{un}}\mathbf{1}_{\text{un}}(ij) + \sum_{k=1}^{5}\beta_k\mathbf{1}_{\text{map}k}(ij), \tag{2}$$

where generically $\mathbf{1}_{\text{Q}}(ij)$ denotes the indicator variable of the ijth observation having the characteristic Q. So $\mathbf{1}_{\text{Q}}(ij) = 1$ if $(ij) = Q$ and is 0 otherwise. This ensures that the appropriate parameters appear in the model in the correct place, i.e. a unique set of parameters appear for each combination of map type and metro system for each individual. So if individual i at task j is looking at the automatically-drawn metro map version of Stockholm metro system, then $x'_{ij}\boldsymbol{\beta} = \beta_{auto} + \beta_3$. The parameters β_{auto} and β_{un} measure the effect on the hazard function associated with the automatically-drawn and the undistorted metro map versions, respectively; whereas the β_ks correspond to factor effects associated with the metro system maps used in the experiment.

Considering a set of data, one can learn from the model (1) via its corresponding likelihood function. It is in this function where the censored data comes into the analysis. A latent variable, δ_{ij}, is defined to be equal to one if the ijth duration time corresponds to a correct answer and zero otherwise. Hence, the likelihood function can be written as

$$L = \prod_{\delta_{ij}=1} h(t_{ij}) \prod_{ij} S(t_{ij}). \tag{3}$$

The parameters to be estimated in the model are $\boldsymbol{\beta} = (\beta_{\text{auto}}, \beta_{\text{un}}, \beta_{\text{map1}}, \ldots, \beta_{\text{map5}})$.

We adopt a Bayesian approach to inference [3]. Hence, all relevant parameters —particularly the parameters β_{auto} and β_{un}— in the model are assumed to be random variables and all of them have associated a probability distribution given the observed data (which is computed via the Bayes theorem using a prior probability distribution for these variables and the likelihood function). Translating Hypothesis A and B into probabilistic statements concerning the βs would lead us to asserting the probability of these hypothesis to be true given the data.

Prior specification and model fitting

The prior probability distribution for h_0 is based on a gamma process [4]. The frailties are assumed to be independent and identically gamma distributed with mean 1 and we assigned a non-informative prior distribution for the fixed effect parameters. Since it is not possible to mathematically derive conclusions from the model, being too complex, we use simulation techniques such as Markov chain Monte Carlo, which samples from the posterior distribution, using standard Markov chain theory, and we can derive conclusions from these samples. Hence, model fitting is carried out using Clayton's algorithm [5] which was implemented in R [6]. The reported results were obtained after running a Markov chain Monte Carlo sampler with 500,000 samples after a burn–in period of 200,000 iterations.

Table 9 shows a summary of the conditional distribution function of the parameters involved in the proportional hazard component (2). The means correspond to the Bayes estimate of the corresponding parameter. Notice that in Table 9 the parameters β_{auto} and β_{un} and the other parameters are specified relative to the official published map and Atlanta's metro system (Model 1). The cities with strong influence on the duration time to complete the task were Mexico City, Stockholm and Toronto. Figure 29 shows the posterior estimate and the 95% confidence bands of the baseline survival function. The variablility across subjects is reflected in the variability of the estimated frailties, see Figure 30.

TABLE 9
Summary of posterior distributions: proportional hazard component.

Factor effect	Model 1 Mean	Model 1 (Std. Err.)	Model 2 Mean	Model 2 (Std. Err.)
Automatically-drawn	0.296	(0.114)	0.270	(0.099)
Undistorted	0.056	(0.117)	-	-
1. Bucharest	-0.223	(0.152)	-	-
2. Mexico City	-0.696	(0.163)	-0.704	(0.172)
3. Stockholm	-0.580	(0.169)	-0.596	(0.175)
4. Toronto	-0.866	(0.181)	-0.872	(0.191)
5. Washington	-0.316	(0.156)	-	-

We fitted an alternative model where we consider the official published and the undistorted metro maps in a single category to contrast Hypothesis A and B simultaneously (A & B). Posterior results for the proportional hazard component of this model are summarized in Table 9 under Model 2.

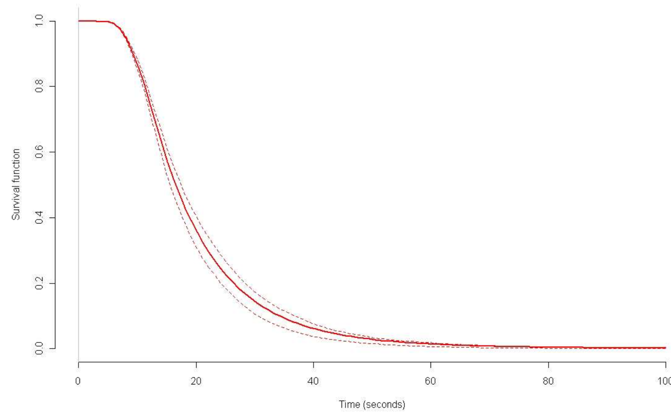


Fig. 29. Posterior estimate of the baseline survival function.

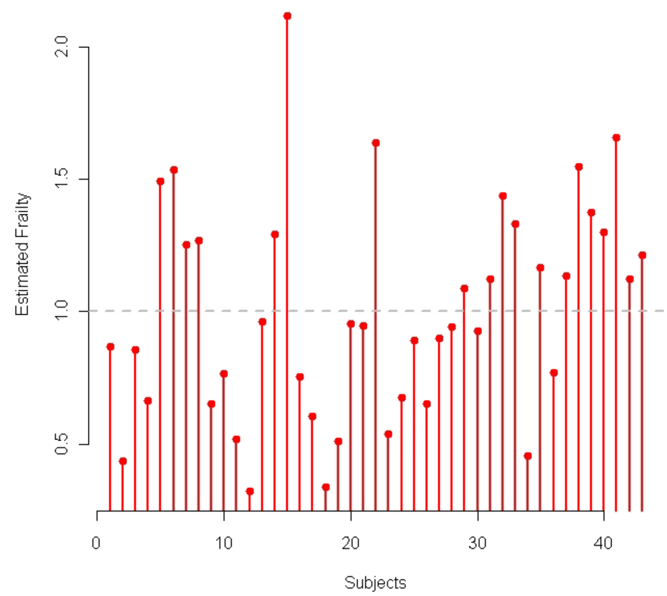


Fig. 30. Posterior frailty estimates.

Contrast of map type effects

Working in terms of Hypothesis A, we can deduce that the automatically-drawn map is preferable to the undistorted map if it is more probable that the time required to find a correct route using the automatically-drawn map is less or equal to the time taken using the published map. That holds when

$$\beta_{\text{auto}} \geq \beta_{\text{un}} \tag{4}$$

and Hypothesis B is supported when

$$\beta_{\text{auto}} \geq 0. \tag{5}$$

From a Bayesian point of view, it is possible to talk about the probability of Hypothesis A (or B) to be true given the observed data. Those quantities can be easily computed using the conditional distribution of β_{auto} and β_{un} given the data. Hence,

$$\Pr(\text{Hyp. A}) = \Pr\left(\beta_{\text{auto}} \geq \beta_{\text{un}} \mid \text{data}\right) \quad (6)$$

$$\Pr(\text{Hyp. B}) = \Pr\left(\beta_{\text{auto}} \geq 0 \mid \text{data}\right). \quad (7)$$

On the other hand, Hypothesis A and B can be simultaneously evaluated considering Model 2 by considering

$$\Pr(\text{Hyp. A \& B}) = \Pr\left(\beta_{\text{auto}} \geq 0 \mid \text{data}\right). \quad (8)$$

Even though it is not possible to compute the above quantities in a closed analytical form, Monte Carlo estimates of these quantities are available (see, [3], for a detailed description).

TABLE 10
Monte Carlo estimates of posterior probabilities: Hypothesis A and B.

Hypothesis	Model 1 Probability	Model 1 (Std. Err.)	Model 2 Probability	Model 2 (Std. Err.)
A	0.986	(0.113)	-	-
B	0.994	(0.076)	-	-
A & B	-	-	0.994	(0.105)

A decision rule in favour of Hypothesis A (or B or A & B) again its alternative would be to have the probability (6) (or (7) or (8)) greater than 0.5. Table 10 shows the Monte Carlo estimates of those probabilities using the two different models. There we can notice some evidence to validate Hypothesis A and B.

C.2 Preference data

In Table 11 we show the contingency table of subjects' preferences across map versions by the map type recorded during the part (iv) of the experiment. From that table, there is a clear preference for the automatically-drawn map, with the exception of Stockholm. For each metro map system there are six possible rankings that a subject can assign to the three map versions, which can be enumerated as follows starting with the most preferable: (a) Auto-Un-Pub, (b) Auto-Pub-Un, (c) Un-Pub-Auto, (d) Un-Auto-Pub, (e) Pub-Auto-Un, (f) Pub-Un-Auto.

Notice that information regarding Hypothesis C and D is summarized in (a) and (b) events. Therefore, contrasting Hypothesis C and D simultaneously (C & D) would be equivalent to contrasting the occurrence of (a) or (b). To assert Hypothesis C & D we have collapsed the six possible rankings into a binary category. The data is analysed using a logistic mixed linear model [7], which is commonly used for such data sets. The model consists of a categorical variable, k, enumerating the two possible outcomes after comparing the three map versions, with $k = 1$ if the automatically-drawn metro map is the most preferable version and $k = 0$ otherwise. Let us denote by y_{ij} the resulting category of the ith subject after ranking the jth map. We assume that the probability of getting either of the possible outcomes is defined as

$$\Pr(y_{ij} = k) = \pi(\eta_i + \boldsymbol{z}_j'\boldsymbol{\phi}), \quad (9)$$

where $\pi(\cdot)$ denotes the logistic function, η_i denotes a random effect attached to subject i, and $\boldsymbol{z}_{ij}'\boldsymbol{\phi}$ denotes a fixed effect associated with characteristics of the jth map, which is defined as

$$\boldsymbol{z}_{ij}'\boldsymbol{\phi} = \phi_0 + \sum_{k=1}^{5} \phi_k \mathbf{1}_{\text{map}k}(ij). \quad (10)$$

We assumed that the η_is are independent and identically Gaussian distributed with mean 0 and a common precision parameter λ. This model, like the one used to analyse the duration data, takes into account map fixed effects and random effects due to repeated observations within each subject.

Prior specification and model fitting

We fitted a semi-parametric version of the logistic regression model expressed as a mixture of Dirichlet processes [8]. We placed a vague prior distribution on the regression parameters. The model was fitted using the `DPpackage` [9] developed in `R`. Results are reported with 500,000 Markov chain Monte Carlo samples recorded after a burn–in period of 100,000 iterations. Table 12 summarizes the posterior distribution for the regression component (10). In Figure 31 we show the estimated effects for subjects.

TABLE 12
Summary of posterior distributions: logistic regression component.

Factor effect	Mean	(Std. Err.)
0. Atlanta	-0.054	(0.366)
1. Bucharest	0.941	(0.493)
2. Mexico City	3.709	(0.887)
3. Stockholm	-2.387	(0.643)
4. Toronto	0.339	(0.473)
5. Washington	1.501	(0.526)

Map version preferences

Notice that Hypothesis C & D are supported when

$$\boldsymbol{z}_{ij}'\boldsymbol{\phi} > 0, \quad (11)$$

for a certain map. Hence, the probability of Hypothesis C & D turns to be

$$\Pr(\text{Hyp. C \& D}) = \Pr\left(\boldsymbol{z}_{ij}'\boldsymbol{\phi} > 0 | \text{data, map}\right). \quad (12)$$

Table 13 shows the Monte Carlo estimates for the probability of Hypothesis C & D by map without considering random effects. In addition, the posterior predictive estimate of the automatically-drawn metro map to be the most preferable map is 0.597 and its posterior 95% confidence interval is $(0.474,\ 0.865)$.

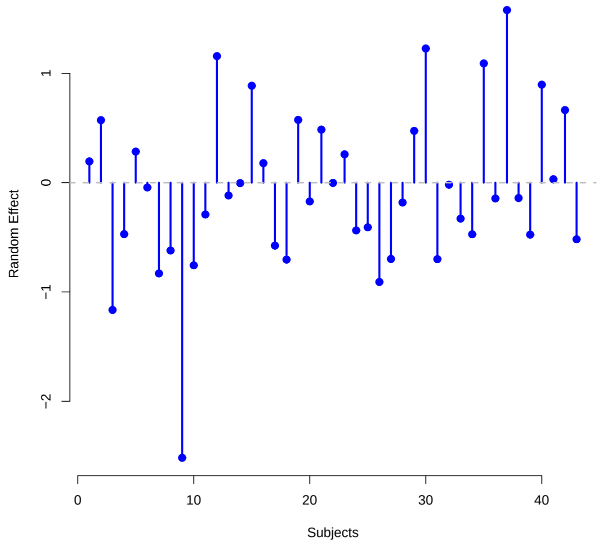


Fig. 31. Posterior estimates for the subjects' effects.

TABLE 13
Monte Carlo estimates of posterior probabilities: Hypothesis C and D.

System	Probability	(Std. Err.)
0. Atlanta	0.439	(0.496)
1. Bucharest	0.990	(0.011)
2. Mexico City	0.993	(0.007)
3. Stockholm	0.008	(0.003)
4. Toronto	0.778	(0.414)
5. Washington	0.995	(0.009)

A decision rule in favour of Hypothesis C & D would be to have the probabilities shown in the above table to be greater than 0.5. Apart from Atlanta and Stockholm, we can conclude that the automatically-drawn metro system was the most preferable version. The cities where the automatically-drawn maps were most preferable are Bucharest, Mexico City, Toronto and Washington.

References

[1] D. Cox, "Regression models and life-tables (with discussion)," *J. the Royal Statistical Society, Series B,*, vol. 34, pp. 187–220, 1972.
[2] D. Clayton and J. Cruzick, "Multivariate generalization of the proportional hazards model (with discussion)," *J. the Royal Statistical Society, Series A,*, vol. 148, pp. 82–117, 1985.
[3] A. Gelman, J. Carlin, H. Stern, and D. Rubin, *Bayesian Data Analysis*. London: Chapman & Hall, 2004.
[4] J. Kalbfleisch, "Non-parametric bayesian analysis of survival time data," *J. the Royal Statistical Society, Series B,*, vol. 40, pp. 214–221, 1978.
[5] D. Clayton, "A monte carlo method for bayesian inference in frailty models," *Biometrics*, vol. 47, pp. 467–485, 1991.
[6] "Statistical analysis: an introduction using R/R basics," http://en.wikibooks.org /wiki/Statistical_Analysis: an_Introduction _using_R/Statistics_and_R, 26 Nov. 2008.
[7] C. McCulloch and S. Searle, *Generalized, Linear, and Mixed Models*. New York: John Wiley & Sons, 2001.
[8] S. Mukhopadhyay and A. Gelfand, "Dirichlet process mixed generalized linear models," *J. the American Statistical Association*, vol. 92(438), pp. 633–639, 1997.
[9] A. Jara, "The `dppackage` package: Bayesian nonparametric and semiparametric analysis, version 1.0-5," http://www2.udec.cl/ ajarav/software.html, 26 Nov. 2008.

TABLE 11
Contingency table of subjects' preferences by map and map type.

Map	*Preference ratings*								
	Most			*Medium*			*Least*		
	Auto	Un	Pub	Auto	Un	Pub	Auto	Un	Pub
0. Atlanta	21	10	12	12	10	21	10	23	10
1. Bucharest	29	9	5	6	22	15	8	12	23
2. Mexico City	41	2	0	1	35	7	1	6	36
3. Stockholm	5	2	36	28	9	6	10	32	1
4. Toronto	24	5	14	17	20	6	2	18	23
5. Washington	33	0	10	9	3	31	1	40	2