

# Proximity, Communities, and Attributes in Social Network Visualisation

Helen C. Purchase  
School of Computing Science  
University of Glasgow  
Glasgow, UK  
helen.purchase@glasgow.ac.uk

Nathan Stirling  
School of Computing Science  
University of Glasgow  
Glasgow, UK  
2254880S@student.gla.ac.uk

Daniel Archambault  
Department of Computer Science  
Swansea University  
Swansea, UK  
D.W.Archambault@swansea.ac.uk

**Abstract**— The identification of groups in social networks drawn as graphs is an important task for social scientists who wish to know how a population divides with respect to relationships or attributes. Community detection algorithms identify communities (groups) in social networks by finding clusters in the graph: that is, sets of people (nodes) where the relationships (edges) between them are more numerous than their relationships with other nodes. This approach to determining communities is naturally based on the underlying structure of the network, rather than on attributes associated with nodes. In this paper, we report on an experiment that (a) compares the effectiveness of several force-directed graph layout algorithms for visually identifying communities, and (b) investigates their usefulness when group membership is based not on structure, but on attributes associated with the people in the network. We find algorithms that clearly separate communities with large distances to be most effective, while using colour to represent community membership is more successful than reliance on structural layout.

**Keywords**— *social networks, community detection, network layout, proximity*

## INTRODUCTION

Community finding is one of the most studied problems in social network analysis. A community typically relates to a cluster in a graph: that is, a group of nodes with many links between members of the group, and few links exterior to that group [1,2]. Several algorithms have been written to automatically detect clusters (and therefore community structures) in networks [3,4,5], with experiments having been conducted to test the strengths of these approaches. Some experiments [6] compared several community detection algorithms with embedded ground truth [7], while others [8] asked human participants to label their communities in their own Facebook network, comparing the performance of community finding algorithms to human-centred ground truth. In all cases, the purpose of these prior experiments was to test the ability of community finding algorithms in recovering the clustering structure embedded in social networks, rather than testing the effectiveness of visual encodings in enhancing human perception of community structure.

Proximity (in the sense of placement of objects in a visualisation) is one of the strongest channels for encoding information visually [9]. Proximity is a well-known Gestalt principle that states that putting objects close together means that they are seen as a group. If nodes are close to each other in the drawing of a graph, users reading the visualisation tend to believe that they are close to each other (in a shortest path sense) even though this may not be the case [10]. Force-directed algorithms, by their nature, explicitly place nodes that are more likely to share a community relationship closer together in the drawing [11,12]. Although this effect is well

known and algorithms have been written to optimise for it, no studies have been run to determine how well these visualisations depict attribute and community relationships along with the structure of the graph.

In this paper, we present a study that explores how well such force-directed approaches can depict community and attribute relationships. We used generated social networks with embedded ground truth [6], thus investigating graphs with a known number of communities based on structural properties. Our six algorithms use different force-directed approaches, allowing us to tune the extent to which attribute or community membership influences the position of nodes. We then ran a user study to determine how accurately a human observer can judge the number of communities in the networks, comparing the effectiveness of these algorithms.

## BACKGROUND

Our literature review covers relevant aspects of the following research topics: pre-attentive visual perception, graph drawing algorithms, social networks, community detection algorithms, and related experimental work..

### A. Visual Grouping

Visual “pop out” is the phenomenon whereby items in a visual scene pre-attentively stand out, and are noticed by a viewer spontaneously [13,14]. Particular visual attributes of objects (e.g. colour, shape, orientation) cause them to pop-out “immediately without effort” [15] when shown amongst a set of objects with different attribute values. Pop-out contrasts with ‘serial search’, where viewers are required to explicitly look at all the objects in a visual scene to locate targets, usually because they need to identify items based on more than one attribute (e.g. red triangles) [13,16]. One of the consequences of this phenomenon is that pre-attentive features can be used to highlight group membership of objects (see Fig. 1a).

The Gestalt grouping principles are based on the view that the perception of an object cannot be reduced to the sum of the perception of its parts [17,18], and that a viewer naturally perceives groups within a scene of objects. These laws are more general than simply defining those visual attributes that ‘pop-out’. They highlight the patterns that we see in visual scenes; for example, how background is distinguished from foreground, that items that look similar or are symmetric to each other appear to be part of a group, and that open shapes are perceived as closed [19].

In this paper, we focus on the pre-attentive visual feature of colour and the Gestalt principle of proximity (that is, items in close Euclidean proximity to each other are perceived as a group, see Fig. 1b), in the context of community membership of social networks, where these networks are depicted as graph drawings.

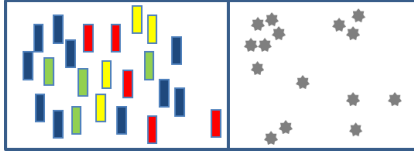


Fig.1a. Colour highlights groups (pop-out): we see the objects of the same colour as belonging to the same group. Fig.1b. Proximity highlights groups (Gestalt): we see objects that are close together as belonging to the same group.

Prior research considering Gestalt principles in relation to graph drawing does not typically consider colour, and is dispersed across the literature. Bennet et al. [20] review the graph drawing aesthetics literature with reference to Norman’s stages of perception [21], including Gestalt theories in the visceral stage, but excluding colour. Marriott et al. [22] investigated features of small graph drawings that make them memorable, considering symmetry, continuity, orientation and proximity, but not colour. They find that drawings with symmetry and continuity are amongst those most readily recalled. In software engineering, Wong and Sun evaluate three UML tools with respect to Gestalt-based criteria [23]; Lemon et al [24] show that similarity, proximity and continuity affect comprehension of complex software diagrams. Neither consider colour. Nesbitt & Freidrich [25] use colour as part of the ‘Law of Similarity’ in relation to graphs evolving over time, not with respect to group identification, but with providing “landmarks” to support maintaining a consistent mental map over time.

### B. Social Networks

Social networks that encode relationships between people (where nodes are individuals and edges are social ties between individuals) are often multivariate networks [26] with a range of different attributes: gender, age, post-code, etc. In graph visualisation these attributes can be depicted using a variety of visual channels: colour, shape, texture etc. (see Fig. 2a).

### C. Graph Drawing Algorithms and Social Networks

Graph drawing algorithms are typically based on several layout principles (minimising edge crossings, highlighting symmetric sub-graphs where possible; maximising the average angle of edges incident to a node etc.) [27]. These algorithms can be applied to graphs that relate to a real domain (e.g. UML diagrams, biological networks), as well as graphs where the nodes and edges bear no relation to any real objects. Indeed, many such graph layout algorithms were initially designed for the latter domain-free case.

Some graph layout algorithms are designed to optimise the visual form of the drawing so as to reveal clusters in the network; that is, sub-graphs that have a high number of edges between the nodes within the sub-graph and only a few edges connected to members outside the sub-graph – also known as communities [1,2]. Energy-based graph layout algorithms are the most common of these; Noack [12] distinguishes between their two components as “an energy model which specifies what layouts to compute, and an energy minimization algorithm which specifies how to compute these layouts”. He defines two novel energy models (the LinLog models [11,12]) which were explicitly designed to depict communities in the drawing by encouraging the placement of nodes within a cluster closer together. Another approach uses ‘virtual springs’ and ‘virtual nodes’ to

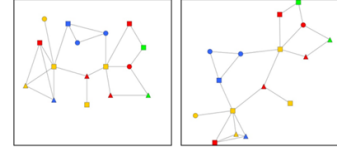


Fig 2a. Example multivariate social network. Edges represent friendships, colours show membership of different committees, shapes indicate one of three age ranges. Fig. 2b. The same network laid out with a force directed algorithm: the structural communities highlighted in the graph layout (friendships) are not the communities represented by the attribute values of the nodes (committees, age).

explicitly pull members of a cluster in towards the centre of that cluster (where the virtual node is placed), thus bringing community members closer to each other [28].

However, these approaches do not provide a way to explicitly adapt the placement of the community nodes or base their position on other attribute values. In this paper, we investigate several simple approaches that are better able to tune the forces that place nodes of the same community membership or attribute value close together.

### D. Communities in Social Networks

Community membership might be explicitly associated with the nodes in a social network as one of their attributes. For example, each member in the network supports one football team in the league, or lives in one postcode in the country, or regularly attends one of many church denominations. Communities may be overlapping or non-overlapping; in the latter case, an individual may belong to more than one community – for example, by regularly attending more than one church. In this paper, we consider strict partitions, that is, nodes belong only to one community.

Where community membership is not explicitly associated with each node as one of its attributes, community membership can be automatically and algorithmically detected using community detection algorithms [3,4,5]. These approaches are based on the principle that the community structure is related to the graph structure – there is a higher likelihood that members of a community have structural relationships (ie. edges) between them than between the members of different communities. Thus, clusters in a graph would be an indication of membership in a community.

This assumption, of course, relies on the structural relationships in the network being in some way related to the community attribute. This assumption is more likely to be valid if, for example, the community attribute is membership of a particular church and the edge relationships are familial, than if the community attribute is membership of a particular church and the edge relationships are co-workers (See Fig 2b).

Lancichinetti and Fortunato [6] performed a large experiment investigating prominent community finding algorithms in the literature. This experiment generated graphs of realistic structure and embedded ground truth using their LFR benchmark approach [7]. A range of community finding algorithms was run on these graphs and their ability to reliably recover the embedded ground truth assessed [29]. The study found that the Infomap community detection algorithm [5] was best able to recover the embedded community structure imposed by the LFR generation approach.

### E. Experimental Work

The early work of McGrath et al [30] initiated experimental work asking participants to identify communities in social networks drawn as graphs. They showed the same graph laid out in five different ways, and asked participants about the number of ‘subgroups’, prominence, and bridging. They found that graph layout affected the distribution of the number of sub-groups identified – the first time this fact had been empirically proven. In their later article [31] they state in their findings that “the number of perceived groups changes when nodes are spatially clustered to hide or highlight a clique”. In this study, participants were asked to assign each node in a graph (laid out three ways) to one (and only one) group. The authors looked at both correct and incorrect group assignment, concluding that spatial clustering affects viewers’ perceptions of groups, and suggesting that nodes connected with edges should be placed near to each other

Lee and Archambault [8] tested how well community detection algorithms recovered known community structures present in Facebook networks. Participants downloaded their own Facebook network and visualised it as a node-link diagram drawn with a standard force-directed algorithm. They were asked to indicate their personal social communities in the network using a lasso selection tool. These communities were recorded and compared to the communities found by these algorithms through normalised mutual information. The study found that Infomap [5] provided the result that best reflected how a human considers communities in their own social network.

Huang et al [32] asked participants to determine how many groups there were in a network (amongst other tasks), having created their graph drawings specifically to conform to five sociogram drawing conventions. They concluded that the layout that “separat[ed] different groups and plac[ed] nodes in the same group close to each other” produced higher accuracy in the group identification tasks, and that this layout was preferred above the others.

As part of a larger study that compared user generated with automatic graph layouts, Dwyer et al [33] asked participants to first select the graph drawing that would best support the task of identifying a six-person clique in a social network (of which there was always only one), before performing the task itself. They found that the highly orthogonal (grid-like) layout produced high accuracy, and that participants focussed on ‘tight clustering, and the shape and symmetry of the clique.’

In all these cases, the community principle used is a structural one; that is, ‘community’ is defined by the presence of several edges between the nodes that belong to the community.

#### RESEARCH AIM

We aim to determine whether explicitly including proximity as a graph layout feature within the design of a layout algorithm when depicting a social network assists in community identification. We have proposed some simple amendments to a common graph layout algorithm that will highlight proximity, and aim to see whether by doing so communities are easy to identify. We also include colour as one method of highlighting community membership, so as to also consider a prominent pre-attentive feature. In our

experiment, participants were asked to count the number of communities in several graph drawings. We wish to determine which layout algorithm better supports identification of communities, and whether this differs according to the size of the graph.

#### A. The Graphs

Many network community finding algorithms could have been used to detect communities in a set of given experimental graphs, but the results of these approaches vary, and may give different answers. Since we require ‘ground truth’ (ie: we need to know exactly which community each node is associated with), we used the LFR benchmark [7]. This had the advantage in that it generated our experimental graphs (in accordance with given parameters) as well as providing community membership ground truth for each of them. In addition, LFR generates graphs that have similar statistical properties to social networks in the real world. The graphs were created using LFR with respect to the following parameters:

- number of nodes: 30, 45, 60, 75.
- minimum number of nodes per community: 5
- number of communities: 3, 4, 5, 6<sup>1</sup>
- mean degree: 4
- maximum degree: 20
- muw (the mixing parameter for the weights): 0.1 (as used in Olsen [34])

All other parameters had default value. There was no requirement that the graphs be connected, but we did insist that non-overlapping communities be defined.

For each graph, we then created a copy, and randomly changed community node membership, keeping the same number and the same size of communities per graph. Doing this provided experimental stimuli for which we know that the communities are not necessarily in close structural proximity – in this latter case, community membership is thus indicated by the attributes of the nodes (AT) rather than by the graph structure (ST), as defined in the original graphs.

This gave us two sets of 16 graphs that include information about non-overlapping community membership for each node: a total of 32 experimental graphs. We use the following notation: [nodes]-[communities]-[principle]. For example, 30-4-st refers to the graph of 30 nodes and 4 communities, where community membership is determined by graph structure; 75-3-at refers to the graph of 75 nodes and 3 communities, where community membership is determined by node attribute.

#### B. The Layout Algorithms

We take as our starting point a common force-directed graph layout algorithm in the knowledge that such algorithms already make an effort to depict nodes that are highly connected in close proximity. We then adapted the algorithm to create versions where community is further

---

<sup>1</sup> It proved impossible to create a graph with 30 nodes and 6 communities of size 5 using LFR. We adapted a 30 node-5 community graph (moving 5 edges, adding 2 edges) to create a 30 node graph that the Louvain Community Detection Algorithm [4] confirmed as having 6 communities.

emphasised by creating additional attractive and repulsive forces in the programming code. This approach is clearly based on the results of the experimental work of McGrath et al [30] and Huang et al [32], which indicate the importance of structural proximity to visual community detection. We also included a version of the algorithm that colours each node according to its known community membership (with no other changes made to the layout algorithm). The algorithms are as follows (see Fig 3):

- **F&R:** This is the ‘control’ algorithm: the force-directed Fruchterman & Reingold algorithm [31] with attractive forces between pairs of nodes connected by an edge and repulsive forces between all pairs of nodes. All nodes are the same colour.
- **COL:** This is the F&R algorithm, with the nodes in each community given a different distinct colour.
- **ATT:** This is the ‘attractive forces’ algorithm. Additional attractive forces are added to the graph between all nodes in the same community, even if they already have an edge (and therefore an attractive force) between them. The extended graph is then laid out using the F&R algorithm. All nodes are the same colour.
- **REP:** This is the ‘repulsive forces’ algorithm. Additional repulsive forces are added to the graph between nodes in different communities, in addition to the repulsive force that already exists between them. The extended graph is then laid out using the F&R algorithm. All nodes are the same colour.
- **A&R:** Both the attractive forces of ATT and the repulsive forces of REP are added to the graph; the extended graph is then laid out using the F&R algorithm. All nodes are the same colour
- **NOD:** An invisible ‘community-node’ is created for each community, and attractive forces are created that link each node to its own invisible community node. The extended graph is then laid out using the F&R algorithm, treating this invisible node in the same way as all other nodes. This is similar to the work of Eades and Huang [28], although their method deliberately placed the virtual node in the middle of the rectangle formed by the 2D extent of the community. All nodes are the same colour.

In all cases (including the original F&R algorithm), the existing and additional forces were all characterised by exactly the same default parameter values to ensure valid comparison, the only difference being the force direction for attractive and repulsive forces. Keeping the parameters the same allows for valid comparison.

These six algorithms were applied to all 32 graphs, creating 180 graph drawings. Our notation is [alg]-[nodes]-[communities]-[principle]. Thus, A&R-45-3-at refers to the graph with 45 nodes and 3 communities, where communities are identified by node attribute, and the graph layout includes both additional attractive and repulsive forces. Since the visual form of F&R-x-y-st is identical to F&R-x-y-at, the 16 stimuli that matched the latter form were removed, leaving 176 stimuli. Before being presented to the participants, each stimulus was rotated randomly ( $0^\circ, 90^\circ, 180^\circ, 270^\circ$ ), minimising the potential for participants recognising repeated stimuli.

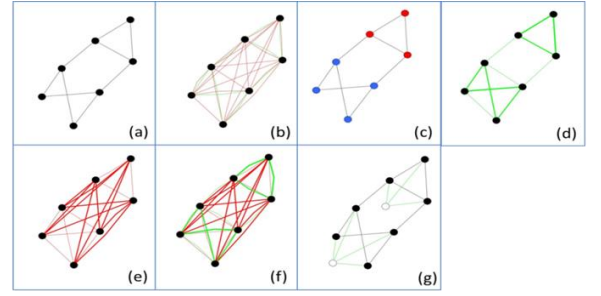


Fig. 3. The forces used in each algorithm. (a) Fruchterman and Reingold layout of the graph. (b) The attractive (green) and repulsive (red) forces used in F&R (F&R). (c) F&R layout showing the two communities by colour (COL). (d) Attractive forces in the ATT algorithm: some attractive forces are reinforced – the repulsive forces in (b) are still there, but not shown. (e) Repulsive forces in the REP algorithm: some repulsive forces are reinforced – the attractive forces in (b) are still there, but not shown. (f) Forces in the combination of ATT and REP (A&R). (g) Two invisible community nodes and attractive forces (NOD).

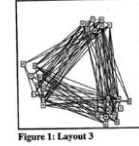


Fig. 4. The nodes in each community are placed in close proximity, despite not forming structural clusters. (Reproduced from McGrath et al, with permission [30]).

### C. Structural vs Attribute Grouping

Each graph has two versions based on two different community principles: a ‘structural’ version (ST) where the membership of each node is determined by the structure of the graph (as dictated by the LFR algorithm), and an ‘attribute’ version (AT) where the membership of each node is unrelated to the structure, having been allocated randomly. There are 96 ST versions and 80 AT versions (since F&R applied to the ST and AT versions are identical.).

Each version will behave differently under the four algorithms that deliberately try to bring communities together, that is, ATT, REP, A&R and NOD. Since the nodes in the ST version are structurally determined, the communities should be structurally accentuated. In contrast, the communities that are unrelated to structure in the AT versions will be brought together, despite not having natural structural proximity. This is not unlike one of the conditions in the McGrath experiment [30], which shows an extreme (hand-drawn) version of communities that are visually spatially distinct, but not structurally connected (see Fig 4).

The coloured nodes (COL) algorithm is, of course, important for AT graphs, since without the colours it would be impossible to determine the community membership of any node – the graph structure will not help. Thus, we have proximity and colour as our visual cues, and structure and attributes as our community membership indicators.

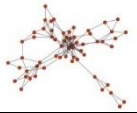
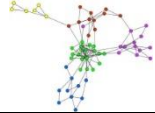


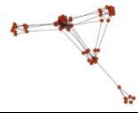

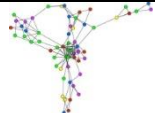




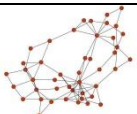
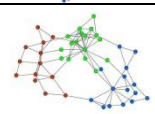




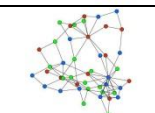




Table 1 shows examples of experimental stimuli; the full set can be seen at [www.dcs.gla.ac.uk/~hpc/ASONAM2020](http://www.dcs.gla.ac.uk/~hpc/ASONAM2020).

### D. Experimental Process

We used a custom-made online system for a within-participants experiment. This allowed us to collect a lot of data quickly and easily, and was suitable as the advantages of lab-based experiments (e.g. training, task complexity etc.) are not relevant for pre-attentive tasks. As a within-participants study, any variation in visual display technology



TABLE 1. EXAMPLE EXPERIMENTAL STIMULI: A 60-NODE GRAPH WITH FIVE COMMUNITIES AND A 45-NODE GRAPH WITH THREE COMMUNITIES.

60-5	F&R	COL	ATT	REP	A&R	NOD
ST 60-5						
AT 60-5	Not Applicable					
ST 45-3						
AT 45-3	Not Applicable					

would have no effect on the results. After reading introductory information and indicating consent, participants were told that “a single circle represents a person in a social network” and “a connection between two circles indicates a friendship between two people”, with appropriate small diagrams to illustrate these concepts. Twelve practice tasks similar in form to the experimental tasks (and for which data were not collected) were presented first, so that the participants could familiarise themselves with the system, and to mitigate against the learning effect. Participants were allowed to take a self-timed break after every 12 trials.

For each of the experimental stimuli, the participants were asked “How many groups of people can you see?” Participants selected one answer from a range of possible answers: 2,3,4,5,6,7, covering all the correct answers (3,4,5,6) with one additional possibility at each extreme. We did not collect response time, since counting six communities will naturally take longer than counting two communities. To move onto the next trial, participants pressed either a button labelled ‘I am confident’ or ‘I am not confident’. We deliberately did not explain what was meant by a ‘group’, since we wanted to participants to form their own ideas, using proximity, colour or structure as they saw fit. At the end of the experiment, demographic information was collected, and we asked participants to tell us what made it easy or hard to identify the groups.

The link to the experiment online system was distributed to friends, family members, associates and colleagues of the authors; as a within-participants experiment, any variation in the nature and expertise of the participants does not affect the results. 42 people took part over 10 days. We removed 3 outliers whose overall accuracy was less than 30%. Our data distribution is such that, in the most extreme case, persistently answering with the same number (e.g. 3) would result in 29% accuracy (thus justifying this 30% cut-off). The majority of the participants (25) were in the age range 18-25; while 17 had background in Mathematics or Computing Science, only three described themselves as experienced in having seen social networks drawn as graphs, with 25 saying that they had ‘never seen’ them like this.

#### DATA AND ANALYSIS

For each trial, we collected the participant’s answer and their confidence, and derived four dependent variables:

- ABV: absolute difference between ‘ground truth’ (GT) number of communities and the answer [0-4]
- SGN: error direction; ie: sign of difference [-1,0,+1]
- ACC: accuracy [0,1]
- CON: confidence: 1 for confident, 0 for not [0,1]

We aggregated these dependent variables (taking the mean) over all variations of the ‘ground truth’ number of communities within each graph size.

We analyse the ST and AT data separately, using SPSS; our research questions are independently addressed for networks where communities are defined by structure (ST) and for those where community membership is independent of structure (AT) (Table 2). Normality tests (skewness, kurtosis and Shapiro-Wilk) revealed that the data is not normally distributed, and so the non-parametric ANOVA alternative was used: Friedman followed by pair-wise comparisons with appropriate p-value adjustments. The median and interquartile ranges for each algorithm for each dependent variable are shown in Table 2 (ST), Table 3 (AT), with the Friedman statistic ( $n=39$ ,  $df=5$ ). Pairwise differences (with adjusted p-values) are in Table 4. As noted above, the F&R algorithm was not relevant for the AT study.

We performed the same analysis on the 30-node graphs and the 75-node graphs separately, to see whether the results differed by graph size. There were no notable significant differences – the separate results mostly matched the results for all graphs when aggregated together, suggesting no effect of graph size.

These results indicate poor performance for some of our new algorithms, with very good performance for others. It appears that the inclusion of additional attractive edges is a positive enhancement, while extra repulsive edges are not useful unless combined with additional attractive edges – although this generalization is not always true when node attributes (rather than structure) indicate community membership. We also note that the drawings produced by the new algorithms are all much more bunched up than the original Fruchterman and Reingold version, with much node occlusion and much smaller bounding boxes. In all cases, the results for the version that used colour are exceptionally good, with these answers provided with high confidence .

TABLE 2. ST RESULTS: THUMBNAIL BOXPLOT OVERVIEWS, MEANS, INTER-QUARTILE RANGES. THE CONDITIONS IN THE BOXPLOTS ARE ORDERED LEFT-TO-RIGHT: F&R, COL, ATT, REP, A&R, NOD; THE RANGE OF THE VERTICAL AXIS IS INDICATED AT THE TOP OF EACH COLUMN.

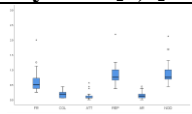
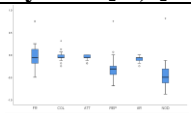
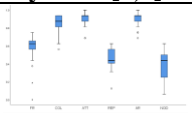
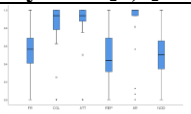
	ABV y-axis: [0,3]	SGN y-axis: [-1,1]	ACC y-axis: [0,1]	CON y-axis: [0,1]
				
<b>F&amp;R</b>	0.50 [0.37]	-0.06 [0.06]	0.63 [0.13]	0.56 [0.31]
<b>COL</b>	0.19 [0.19]	-0.06 [0.06]	<b>0.88</b> [0.13]	0.94 [0.25]
<b>ATT</b>	0.06 [0.07]	-0.06 [0.06]	<b>0.94</b> [0.06]	0.94 [0.12]
<b>REP</b>	0.75 [0.37]	-0.31 [0.19]	0.44 [0.18]	0.44 [0.37]
<b>A&amp;R</b>	0.13 [0.13]	-0.06 [0.07]	<b>0.94</b> [0.06]	1.00 [0.06]
<b>NOD</b>	0.75 [0.31]	-0.50 [0.32]	0.44 [0.25]	0.50 [0.38]
<b>Friedman</b>	164.0, p<0.001	115.6, p<0.001	166.7, p<0.001	133.5, p<0.001

TABLE 3. AT RESULTS: THUMBNAIL BOXPLOT OVERVIEWS, MEANS, INTER-QUARTILE RANGES. THE CONDITIONS IN THE BOXPLOTS ARE ORDERED LEFT-TO-RIGHT: F&R, COL, ATT, REP, A&R, NOD; THE RANGE OF THE VERTICAL AXIS IS INDICATED AT THE TOP OF EACH COLUMN.

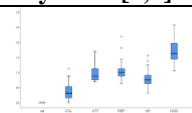
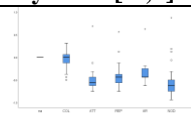
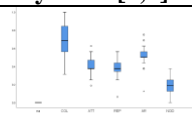
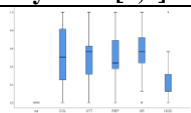
	ABV y-axis: [0,3]	SGN y-axis: [-1,1]	ACC y-axis: [0,1]	CON y-axis: [0,1]
				
<b>COL</b>	0.31 [0.43]	0.00 [0.25]	<b>0.69</b> [0.32]	0.50 [0.63]
<b>ATT</b>	0.88 [0.38]	-0.56 [0.19]	0.38 [0.12]	0.56 [0.32]
<b>REP</b>	1.05 [0.25]	-0.44 [0.19]	0.38 [0.13]	0.44 [0.37]
<b>A&amp;R</b>	0.75 [0.31]	-0.44 [0.19]	<b>0.50</b> [0.06]	0.56 [0.31]
<b>NOD</b>	1.63 [0.56]	-0.63 [0.25]	0.19 [0.12]	0.13 [0.18]
<b>Friedman</b>	129.3, p<0.001	86.3, p<0.001	129.4, p<0.001	63.1, p<0.001

TABLE 4. ADJUSTED PAIRWISE P-VALUES BETWEEN THE LAYOUT CONDITIONS. ONLY SIGNIFICANT DIFFERENCES ARE SHOWN.

	ST				AT			
	ABV	SGN	ACC	CON	ABV	SGN	ACC	CON
F&R/ COL	0.001		<0.001	<0.001	n/a	n/a	n/a	n/a
F&R/ ATT	<0.001		<0.001	<0.001	n/a	n/a	n/a	n/a
F&R/ REP		<0.001			n/a	n/a	n/a	n/a
F&R/ A&R	<0.001		<0.001	<0.001	n/a	n/a	n/a	n/a
F&R/ NOD	0.025	<0.001	0.015		n/a	n/a	n/a	n/a
COL/ ATT					<0.001	<0.001	<0.001	
COL/ REP	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	
COL/ A&R						0.001		
COL/ NOD	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
ATT/ REP	<0.001	<0.001	<0.001	<0.001				
ATT/ A&R						0.004	0.016	
ATT/ NOD	<0.001	<0.001	<0.001	<0.001	<0.001		<0.001	<0.001
REP/ A&R	<0.001	<0.001	<0.001	<0.001	0.003		0.003	
REP/ NOD					<0.001	0.016	<0.001	<0.001
A&R/ NOD	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

## DISCUSSION

In the structural membership (ST) case, the results suggest [ATT, COL, A&R] as the best, [REP, NOD] as the worst, and [F&R] in the middle; the confidence results mirror the accuracy data. The algorithms that performed well (ATT, A&R) clearly depict large inter-community distances. These two algorithms add additional intra-community attractive forces between groups of nodes within each community where there is already a high proportion of attractive forces (as dictated by edge connectivity). Given that there are few attractive edges between communities, this high concentration of attractive forces within communities overwhelms the effect of repulsive forces between them, providing good separation between individual groups. The only repulsive forces with any noticeable effect are those between communities, thus increasing this separation.

The two algorithms that performed worse (NOD, REP) both separate communities, but emphasise intra-community proximity at the expense of inter-community distance. In REP, repulsive forces are added between nodes of different communities; applying these forces to all nodes in a single community has the effect of pushing communities apart; there is little intra-community attraction to bring the nodes together. Thus, the spacing between nodes within a community and between communities becomes uniform, causing participants to visually merge them into one community.

Finally, the NOD condition adds a large number of attractive forces pulling nodes in communities together, but these overwhelm the existing repulsive forces, so communities tend to occlude each other.

Neither F&R nor COL emphasise community distance as clearly as ATT and A&R – they have a spatial balance between connectivity and distance; it is not surprising F&R performed worse than ATT and A&R. COL outperforms F&R because of additional visual coding – even though we did not suggest to our participants that colour relates to group membership.

These results suggest that adaptations to algorithms that clearly emphasise the distinction between communities through increased inter-community distances (thereby reducing the intra-community node distances) is preferable to a neutral algorithm that does not attempt to emphasise structure – unless colour is used. F&R performs in the middle, suggesting it provides a good compromise between community representation and other readability tasks.

For non-structural membership (AT), the results suggest [COL, A&R] as the best, [NOD] as the worst, [ATT, REP] in the middle, and equal confidence for all conditions apart from NOD (about which participants were very unconfident). It is clear that the A&R algorithm's use of both additional repulsive and attractive forces means that each community's nodes are depicted in close proximity, almost independently of any of the existing forces in the basic force-directed algorithm. It maintains intra-community forces (even if there are no edges between the nodes) while also emphasising inter-community distances (even if there are edges between nodes from different communities). The output of this algorithm is similar to the hand-drawn network shown in Fig. 4, where a person's membership of a community (as indicated by attribute) is prioritised over her relationships with others (as indicated by edges).

ATT and REP each only add one type of force (unlike A&R) and so are unable to balance the need to pull intra-community nodes together at the same time as pushing inter-community nodes apart.

NOD's performance is very poor indeed, and all participants were less confident with this condition than others. Here a large number of attractive forces are being added, often between nodes that did not originally have an edge (and therefore an attractive force) between them. This results in the sort of 'hairball' that usually arises from application of a force-directed algorithm to a large, dense graph – and in this case, the communities occlude each other to the extent that they are indistinguishable.

COL again performs well; it appears that for this condition, participants simply counted the colours, ignoring structure. This is surprising since for all other conditions colour was clearly irrelevant, and node proximity was the key to finding the answer. Indeed, the high performance of the COL condition suggests that any tweak made to a layout algorithm to emphasise non-structural community membership using proximity is not as effective as simply using a neutral layout and colour.

In all cases (ST and AT), the mean estimation error was negative: participants tended to underestimate rather than overestimate the number of communities. This is not surprising, since in no conditions are the intra-community distances (that is, proximity between nodes in the same community) dispersed, visually breaking up a single community into multiple communities; thus, it is unlikely participants would see two clusters where there is only one.

When focusing on comparing the F&R and COL conditions, we calculated the correlation co-efficient between the answers given to both (for each stimulus and participant), regardless of the accuracy of the answers. We expected a high correlation for the ST data (the colour in COL reinforcing the response given to F&R), and a lower one for the AT data (F&R responses determined by structure; COL responses determined by colour), but found that they were not very different from each other (ST correlation = 0.591, AT correlation = 0.589, both with  $p < 0.001$ ). It appears that the COL condition performed well for both the ST and AT stimuli regardless of the underlying structure imposed by the F&R algorithm that it was based on.

The qualitative data revealed that, when identifying groups, participants relied on 'clusters' and 'clumps' (16), 'proximity' and 'distance' (15), and colour (16). Mention was also made of 'edge density' or 'number of connections' (8), and one participant "ignored colour". 26 participants mentioned colours when asked what made it easier to identify the groups; 8 people specifically said that 'random colours' made the task more difficult. The feature that most participants said confused them was 'dense clusters' or 'overlapping of groups'.

#### A. Subjective data

In an attempt to further distinguish between the utility of the A&R and COL algorithms (which performed equally well in the objective measures), we conducted a supplementary experiment with 25 new participants, collecting subjective opinion data. Participants were presented with two layouts of each graph (A&R and COL) and asked "Which social network best represents communities?" The graphs were separated according to ST and AT versions, the pairs were matched with respect to size and number of communities, each pair was presented twice (alternating left with right) and shown in random order, and each graph drawing was randomly rotated by a multiple of 90 degrees before being displayed.

Using a Binomial test of frequencies, we found that the A&R algorithm was chosen as best for representing communities for both structural (ST) and attribute (AT) communities, over all the graphs ( $p < 0.001$  for ST,  $p < 0.01$  for AT). When separated by graph size and number of communities, we found only one significant result for the ST graphs (75-3, in favour of A&R), but eleven significant results for the AT graphs (five in favour of COL (30-4, 30-5, 45-5, 45-6, 60-6) and six in favour of A&R (45-3, 45-4, 60-3, 60-4, 75-3, 75-4), all  $p < 0.01$ ). These results suggest that, in general, the Gestalt structural grouping properties of A&R are preferred to the pre-attentive colour properties of COL.<sup>2</sup>

#### CONCLUSION

Our adaptations to the common Fruchterman & Reingold algorithm are simple, both in concept and implementation, and were clearly designed to decrease the distance between nodes in the same community and to increase the distance between nodes in different communities, using additional attractive and repulsive forces. We find that some are more successful than others, and that their utility differs depending on whether 'communities' in a social network are defined by

<sup>2</sup> A result that will, of course, also apply if colour-blindness or cultural colour associations were to be taken into account – two issues that are beyond the scope of this current study.

structure (relationships between people) or attributes (characteristics of people).

If community membership is defined by *structure*, algorithms that include forces that deliberately bring nodes in the same community in close proximity are more successful than those that simply attempt to push members of disparate communities apart. Using colour for depicting community membership is also effective in this case.

If communities are defined by *attributes*, colour is by far the most effective means of portraying membership, and the results suggest that it might be better to abandon any attempt to manipulate the layout algorithm, and to simply use colour coding if possible. Otherwise, the combination of pulling related nodes together at the same time as pushing unrelated ones away also shows promise.

In both cases, gathering all the nodes in a community around an invisible anchor node produced poor performance, since such an algorithm leads to overlapping communities.

The generalizability of these results is, of course, limited by the extent of our experimental parameters, in particular the nature and size of the graphs, and the participants. Our use of LFR means that our graphs are comparable in structure to real-world social networks, and our analysis shows no difference in performance between the 30-node and 75-node graphs. Investigating larger graphs could be a further study, although given our results, such a study should likely focus on ATT, REP and A&R only. Investigating algorithms that define groups by the combination of structure and attributes is an obvious next step for this work. While the participants' age range was rather narrow, they were overwhelmingly novices in reading social networks as graphs, and so our results are not tainted by experience.

This study has demonstrated the utility of making simple adaptations to a common force-directed algorithm for the purposes of visual community detection. Our experimental data confirms that the use of both colour and easily perceived Euclidean proximity as visual cues are crucial for identifying communities, with colour proving more prominent.

#### ACKNOWLEDGMENTS

Thanks are due to all our participants, and to John Hamer who created essential supporting utilities. Ethical clearance was given by the University of Glasgow (ref 300190021).

#### REFERENCES

- [1] Girvan M, Newman MEJ. Community structure in social and biological networks. *PNAS*. 2002;99(12):7821-6.
- [2] Fortunato S, Castellano C. Community Structure in Graphs. In: Meyers R, editor. *Encyclopedia of Complexity and System Science*: Springer; 2009.
- [3] Newman MEJ, Girvan M. Finding and evaluating community structure in networks. *Phys Rev E*. 2004;69(2):026113.
- [4] Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *J. Statistical Mechanics: Theory and Experiment*. 2008:P10008(10).
- [5] Rosvall M, Bergstrom CT. Maps of random walks on complex networks reveal community structure. *PNAS*. 2008;105(4):1118-23.
- [6] Lancichinetti A, Fortunato S. Community detection algorithms: A comparative analysis. *Physical Review E*. 2009;80:056117.
- [7] Lancichinetti A, Fortunato S, Radicchi F. Benchmark graphs for testing community detection algorithms. *Physical Review E*. 2008;78:046110.

- [8] Lee A, Archambault D. Communities Found by Users -- not Algorithms: Comparing Human and Algorithmically Generated Communities. *CHI: ACM*; 2016. p. 2396-2400.
- [9] Munzner T. *Visualization Analysis and Design*: CRC Press; 2014.
- [10] Fabrikant S, Montello DR. The effect of instructions on distance and similarity judgements in information spatializations. *International Journal of Geographical Information Science*. 2008;22(4):463-78.
- [11] Noack A. An Energy Model for Visual Graph Clustering. In: Liotta G, ed. *Graph Drawing Symposium*: Springer, LNCS 2912; 2003. p. 425-36.
- [12] Noack A. Energy Models for Graph Clustering. *Journal of Graph Algorithms and Applications*. 2007;11(2):453-80.
- [13] Treisman A. Preattentive Processing in Vision. *Computer Vision, Graphics, and Image Processing*. 1985;31:156-77.
- [14] Saarinen J. Localization and Discrimination of "Pop-out" Targets. *Vision Research*. 1996;36(2):313-6.
- [15] Theeuwes J, Lucases MP. An Adaptation-Induced Pop-Out in Visual Search. *Vision Research*. 1993;33(16):2353-7.
- [16] Reingold EM, Behrmann M, Moscovitch M. Patterns of eye movements during parallel and serial visual search tasks. *Canadian Journal of Experimental Psychology*. 1997;51(2):151-64.
- [17] Rock I, Palmer S. The Legacy of Gestalt Psychology. *Scientific American*. 1990;236(6):84-91.
- [18] Palmer SE. Theoretical approaches to vision. In: Palmer SE, editor. *Vision Science: Photons to Phenomenology 1999*. p. 45-92.
- [19] Palmer SE. Goodness, Gestalt, Groups and Garner. In: G. R. Lockhead & J. R. Pomerantz, editors. *The Perception of Structure* American Psychological Association; 1991.
- [20] Bennett C, Ryall J, Spalteholz L, Gooch A. The Aesthetics of Graph Visualisation. In: Cunningham DW, Meyer G, Neumann L, editors. *Computational Aesthetics in Graphics, Visualization, and Imaging 2007*. p. 1-8.
- [21] Norman DA. *Emotional Design: why we love (or hate) everyday things*: Basic books; 2004.
- [22] Marriott K, Purchase HC, Wybrow M, Goncu C. Memorability of Visual Features in Network Diagrams. *IEEE TVCG*. 2012;18(12):2477 - 85.
- [23] Sun D, Wong K. On evaluating the layout of UML class diagrams for program comprehension. 13th Int. Workshop on Program Comprehension: IEEE; 2005. p.317-26.
- [24] Lemon K, Allen EB, Carver JC, Bradshaw GL. An empirical study of the effects of Gestalt principles on diagram understandability, 1st International Symposium On Empirical Software Engineering and Measurement, 2007. p. 156-65.
- [25] Nesbitt K, Friedrich C. Applying Gestalt Principles to Animated Visualisations of Network Data. *Information Visualisation: IEEE*; 2002. p.737 - 43
- [26] Kerren A, Purchase HC, Ward MO, editors. *Multivariate Network Visualization*: Springer Verlag; 2014.
- [27] Coleman MK, Parker DS. Aesthetics-based graph layout for human consumption. *Software Practice & Experience*. 1996;26(12):1415-38.
- [28] Eades P, Huang ML. Navigating Clustered Graphs using Force-Directed Methods. *Journal of Graph Algorithms and Applications*. 2000;4:157-81.
- [29] Danon L, Diaz-Guilera A, Duch J, Arenas A. Comparing community structure identification. *J. Statistical Mechanics: Theory and Experiment*. 2005:P09008.
- [30] McGrath C, Blythe J, Krackhardt D. Seeing Groups in Graph Layouts. *Connections*. 1996;19(2):22-9.
- [31] McGrath C, Blythe J, Krackhardt D. The effect of spatial arrangement on judgments and errors in interpreting graphs. *Social Networks*. 1997;19:223-42.
- [32] Huang W, Hong S, Eades P. Effects of sociogram drawing conventions and edge crossings in social network visualisation. *Journal of Graph Algorithms and Applications*. 2007;11(2):379-429.
- [33] Dwyer T, Lee B, Fisher D, Quinn KI, Isenberg P, Robertson G, et al. A comparison of user-generated and automatic graph layouts. *IEEE TVCG*. 2009;15(6):961-8.
- [34] Olsen MJ. *Community Detection in Large Social Networks*: Norwegian Technical Science University; 2014.
- [35] Fruchterman TMJ, Reingold EM. Graph drawing by force-directed placement. *Software - Practice and Experience*. 1991;21(11):1129-64.