Asymmetric Relations in Longitudinal Social Networks

Ulrik Brandes and Bobo Nick

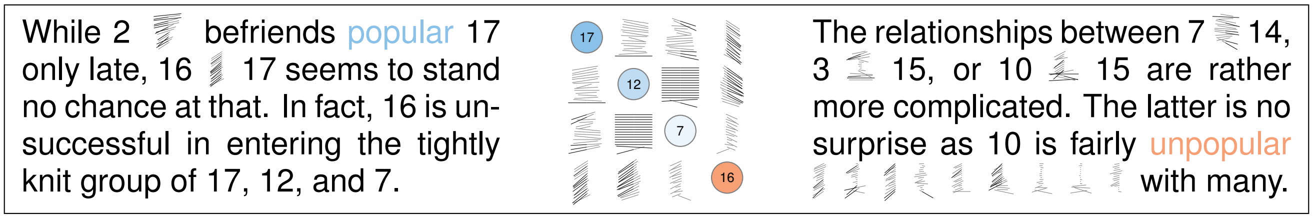


Fig. 1. The emotional seesaws of a liking relationship (extent and balance, dyads to be read bottom to top).

Abstract—In modeling and analysis of longitudinal social networks, visual exploration is used in particular to complement and inform other methods. The most common graphical representations for this purpose appear to be animations and small multiples of intermediate states, depending on the type of media available. We present an alternative approach based on matrix representation of gestaltlines (a combination of Tufte's sparklines with glyphs based on gestalt theory). As a result, we obtain static, compact, yet data-rich diagrams that support specifically the exploration of evolving dyadic relations and persistent group structure, although at the expense of cross-sectional network views and indirect linkages.

Index Terms—Network Visualization, Social Networks, Time Series Data, Visual Knowledge Discovery and Representation, Glyph-based Techniques.

1 Introduction

Social networks are constructs capturing interdependencies among the states and actions of seemingly autonomous social actors [3]. The two principle aims in the study of social networks are understanding their formation (networks as consequences, dependent variables) and their effects (networks as antecedents, explanatory variables) [2]. For empirically observed associations between network structure and actor behavior, however, there are often competing explanations with opposite directions of causality. Longitudinal social network data, i.e., network data over time, are thus crucial to assess whether the social embedding of an actor influenced the actor's actions (social influence), or whether an actor's actions prompted a change of relations (social selection).

Visualization is an essential tool in both the exploration of social networks and the communication of findings [20, 5, 12]. While cross-sectional network data pose many challenges already, longitudinal data increase the level of complexity significantly [35]. Because of this complexity and the multitude of interests that analysts may have, it is unlikely that there are general visualization schemes serving the majority of needs. Instead, social network visualizations should be tailored to the specific type of data and analytical interest at hand.

Here we are interested in such a specific scenario, assuming a network of asymmetric relations and an interest in the evolution of dyadic (i.e., pairwise) relations along with their embedding in the structure at large. This scenario is rather typical for empirical studies of longitudinal social networks.

In the visualization method we propose for this scenario, specifically designed gestaltlines (cf. Section 3.1) – the gestalt-based use of glyphs in sparklines [54] for multivariate sequences – are integrated with a matrix representation of all dyads in the network. The resulting diagrams share the overall strengths and weaknesses of matrix representations [22], but in addition allow for the simultaneous representation of all evolving dyads in a single, static image suitable for print publication.

The remainder is organized as follows. We start with some background on longitudinal social networks and their visualization in the next section. The cornerstone of our approach is a static representation of the evolution of certain types of dyadic data and derived in Section 3. In Section 4, these elements are combined into a representation of entire longitudinal networks. The approach is illustrated by a case study on a well-researched data set presented in Section 5, and initial feedback from domain experts reported in Section 6. We conclude with a brief discussion including directions for future work.

2 Background

In this section, we delineate the scope of our method by spelling out in detail the type of data assumed and by reviewing related work on its visualization.

2.1 Social Network Data

The most common type of data in empirical social network studies consists of a set of actors such as individuals or organizations, and one or more types of relations between them. Examples are friendship networks among pupils and the trading of goods between nations. Relations may be symmetric (as is often assumed in the former example) or asymmetric (as in the latter example). Often, relations are valued, i.e., a numerical value associated with each tie indicates a specific quality of the relation. See [25, 34, 37] for background on social network data and data collection.

We here assume networks of a single valued asymmetric relation. These are modeled by an edge-weighted directed graph $G = (V, E; d)$, consisting of a set of vertices V representing the actors, a set of directed edges $E \subset V \times V \setminus \{(v,v) : v \in V\}$ representing the ties, and a real value $d : E \to [0, M]$ in the range from zero to a maximum value M associated with each edge. We will use $d_{ij} = d((i, j))$ to denote the value associated with edge $(i, j) \in E$, and do not distinguish $d_{ij} = 0$ and $(i, j) \notin E$.

Note that we do not assume actor attributes, because their number and type is much more application-specific. Note also that the majority of empirical studies deals with networks of 10–100 actors.

• *Ulrik Brandes is with the Department of Computer & Information Science at the University of Konstanz, E-Mail: ulrik.brandes@uni-konstanz.de.*
• *Bobo Nick is with the Department of Computer & Information Science at the University of Konstanz, E-Mail: bobo.nick@uni-konstanz.de.*

Manuscript received 31 March 2011; accepted 1 August 2011; posted online 23 October 2011; mailed on 14 October 2011.
For information on obtaining reprints of this article, please send email to: tvcg@computer.org.

2.2 Longitudinal Networks

We have argued in the introduction that longitudinal data is essential in assessing causality. The addition of (forms of) time-variability, however, leads to a combinatorial explosion of data types and problems.

Here, we assume panel data (waves of network observations rather than, e.g., dyadic events) on a fixed set of actors. This is the standard scenario for the form of analysis that is most common in the social sciences and based on stochastic actor-oriented models [46].

The data thus consists of a sequence of weighted directed graphs $G_1 = (V, E_1; d^1), \ldots, G_T = (V, E_T; d^T)$ sharing the same vertex set V, but with potentially varying edge sets and edge weights. Typical values for T are in the range of 2–30 observations, with a bias toward the lower end.

2.3 Network Visualization

Traditional graphical representations of social networks are the common *sociogram* (node-link representations as in Figure 3(d)) and the less common *sociomatrix* (a matrix representation as in Figure 7(a)). See [36] for some of the earliest examples.

The crucial algorithmic challenges for node-link diagrams are placement of nodes and routing of edges. Such layout problems are the main focus in graph drawing [1, 31], where many fundamental techniques have been developed that can be adapted for specific scenarios. Exemplary designs are proposed, e.g., in [7, 6, 43]. A subcategory of node-link diagrams is formed by attribute-based designs as exemplified in PivotGraphs [56].

Similarly, the main algorithmic challenge for matrix representations is the ordering of rows and columns [13]. In combination with specific interaction concepts they have re-gained some popularity in systems for exploring large network data such as [16]. A hybrid approach combining node-link diagrams and matrix representations for dense subgraphs is presented in [29].

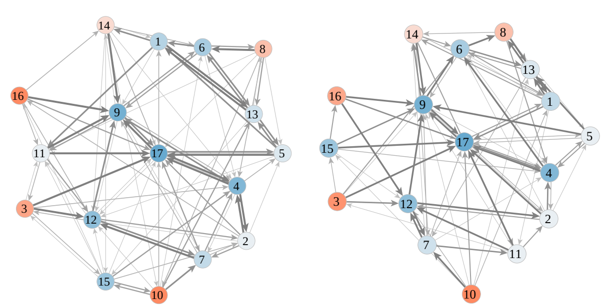
2.4 Time-Varying Network Visualization

A straightforward solution is the use of animation. Since animation is a mapping of empirical time to display time it requires special media. For node-link representations, the layout problem is made more severe by the additional coherence constraints required to ensure smooth transitions between high-quality layouts of $G_1, \ldots, G_T$. A common technique is the anchoring approach of [8, 21], and a review is given in [35].

An alternative to plain animation is interaction, in which time and focus are chosen explicitly as in the matrix-based system of [59]. See also [50] for a systematic treatment of this approach.

If dynamic media are not available or simultaneous cross-time comparisons are of importance, snapshots of an animation are often displayed in small multiples [53]. The example in Figure 3 is based on an aggregation approach, in which all nodes remain in the same position determined from a network aggregated over all time points. Such aggregate networks can be used for representation themselves as in Figure 2. Animation and small-multiples are treated in depth in [4].

We are interested in static representations suitable for traditional publication media, but instead of aggregating observations into a summary network, they will be integrated in a single view without information loss.



(a) first five weeks (b) last five weeks

Fig. 2. Example for small multiples of aggregate summary views on network panel data. Link prominence according to cumulated top 3 friendship nominations in the Newcomb Fraternity Data. Node coloring according to overall popularity; compare gestaltmatrix representation below.

3 Dyad Evolution

As motivated above, we are looking for an intuitive way to communicate entire dyadic evolution clearly and effectively through simple graphical means. In a second step, then, such a representation shall be incorporated into a matrix view, conveying the entirety of dyadic time-series in network panel data in a single diagram to allow for visual exploration of patterns, trends, and outliers.

3.1 Concept: Gestaltlines

The proposed design is a prototypical application of what we termed *gestaltlines*. It is, hence, inspired by a combination of three powerful concepts from information visualization:

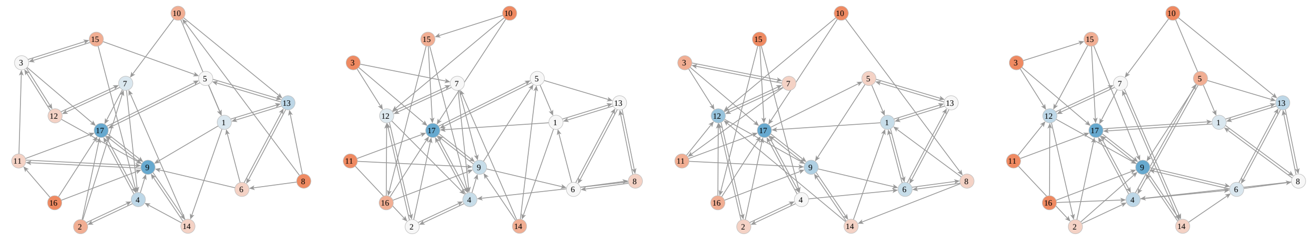
Tufte's Sparklines are "data-intense, design-simple, word-sized graphics" [54]. These *datawords* are conventionally used for univariate time-series data – potentially enriched with additional annotation to convey, e.g., normal ranges and individual data points, such as glucose 6.6 [54, p. 47] – but applicable in other scenarios such as in [24] as well. On simple (high-resolution, colored) printouts, sparkline-like diagrams allow for the display of huge amounts of data within eyespan. The latter is useful especially for exploration, when it is not clear from the start which quantities to compare.

Gestalt theory deals mainly qualitatively with human biases towards perceiving general forms, meaning that "the whole is different from the sum of its parts". According to Wertheimer [57], the *Law of Prägnanz* causes the mind to organize originally disparate visual stimuli into the simplest stable and coherent form (see, e.g., [48]), such as grouping together similarly looking, spatially close, uniformly connected, and symmetrical [ ] < { } > elements, or layering elements according to prominence (law of figure and ground). These principles guide design choices, e.g., in visual screen design [10], human-computer interaction [18], information dashboard design [17] or animated visualizations of network data [39].

Multivariate Glyphs originate from systematic mappings of multiple data attributes onto graphical concepts (see, e.g., [55]). A popular example are star plots [9], in which line segments that radiate out from a common origin represent different data dimensions. Connecting their outer endpoints and filling the center creates a holistic gestalt that is more easily memorized and compared than the data tuple.

Obviously, glyph-based techniques allow to extend classical sparklines to multivariate datawords. Dense, data-rich diagrams, however, do not necessarily lend themselves to visual promotion of any kind of pattern. Therefore, *gestaltlines* shall be combinations of glyphs and sparklines that are directly inspired by gestalt principles: If patterns in the data are bound to invoke gestalt laws, a dataword is capitalizing on the holistic capabilities of human vision by conveying trends, transitions, outliers, or other sequence patterns.

The derivation of corresponding representations for asymmetric relations in longitudinal networks is illustrated below. The following textbook example involving univariate sequences, however, should make the idea obvious: Sedgewick [44] illustrates the procedure of bottom-up mergesort by depicting an initial, an intermediate and the final sorted sequence using sloped lines that indicate numerical values. Their gradually changing slopes prompt us to perceive sorted subsequences as a whole, in which outliers are detected easily.

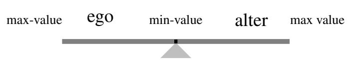


(a) wave 2 (b) wave 6 (c) wave 10 (d) wave 14

Fig. 3. Small multiples of network panel data with uniform outdegree. Red/blue color scale represents lower/higher-than-expected indegree.

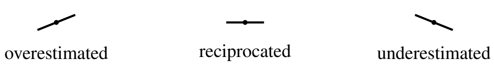
3.2 Basic Design

A simple design, exploiting the impact of gestalt principles and mapping relational/dyadic (ego-alter) data quite naturally to graphical concepts, emanates from the metaphor of a seesaw or scale, respectively. More specifically, take a transparent tube which is subdivided and pivoted in the middle:



Being empty in the beginning it is successively (un)filled from the middle, according to the extent of ratings on a relation

which, of course, may result in imbalance



With this rather minimalistic glyph design, including the additional dimension of time to form a single (letter-like) dataword is now straightforward. Some may find the stacking below reminiscent of a modernistic depiction of a falling tube tumbling left and right according the balance of its filling. This is intentional because it is in line with our basic metaphor. Since human perception is geared toward comparison of progress with respect to some grounded baseline, however, we use bottom-up visualizations of development instead; compare Figure 4.

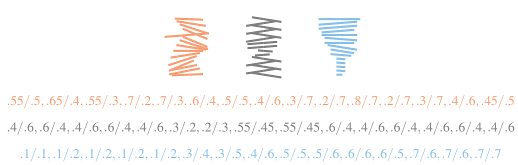


Fig. 4. Basic graphical design (read bottom to top) for any dyadic time-series (read left to right). Typical examples would be ego-alter ratings on a relationship. Patterns, trends, and outliers are clearly visible and interpretable — "Their relationship was an emotional seesaw."

The graphical mapping displays both quantities of interest, the absolute values (extent) and their differences (balance/reciprocity), simultaneously:

- the higher the weight on an edge, the more ink (irrespective of the degree of balance)
- the more imbalance in a dyad, the more prominent the slope (irrespective of the size of weights)

Our design choices are based on intuition and gestalt principles. While a single seesaw intuitively represents the concept of imbalance, their alignment is intended to invoke the gestalt principle of *common fate*, which states that graphical elements of similar orientation are perceived as a whole even when they differ in other attributes such as size. Similarly, motivations to enhance the basic design are addressed in Section 3.3.

The conscious exploitation of gestalt principles appears to add a new twist to the use of glyphs and may contrast the proposed graphical mapping from other designs working also in small-scale such as juxtaposed bar charts mapping time on the vertical axis (Figure 5), rotated horizon graphs [28] or data vases [49].

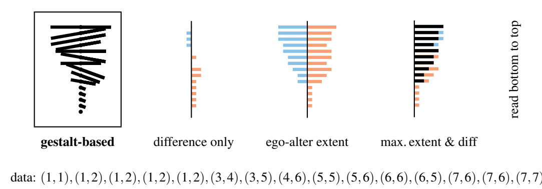


Fig. 5. Gestalt-based representation of dyad evolution compared to conventional times series charts displaying the difference of values, contrasting both values, or juxtaposing maximum extent and difference.

On a second level, the alignment of multiple dyad time series gestalts in a matrix arrangement is expected to support exploration of patterns for egos, alters, and groups of them. More details are provided in Section 4.

3.3 Enhancements

The minimalist nature of the basic glyph design allows for additional enhancements, inspired by gestalt laws.

First of all, the current design leaves color as a degree of freedom to represent additional attributes; cf. [26] for various coloring schemes. For instance, single data values or special data ranges can be highlighted or mitigated, exploiting the law of figure and ground, as exemplified within the case study below. Another particular application in social networks would be the fading out of unconfirmed relations because these may be a critical factor in relational data collection [33]. Alternatively, color could be used to combine and distinguish different ranges of extent and/or balance as suggested by the law of similarity.

We already alluded to claim that human perception subconsciously relates progress with respect to some baseline. One may capitalize on this by explicitly indicating a reference line that incorporates additional (statistical) information at the bottom of each gestalt. We even suggest to use two reference lines where appropriate, e.g., one indicating the expected time-series values in the beginning and one indicating the expected values at the end of the evolution.

Although this is rarely the case in network survey data, we do note that if there are too many points in time, a horizontal representation of evolution might be more appropriate for word-like representations. That is, vertical lines are depicted next to each other, metaphorically indicating balance as "stable" | , or toppling towards ego \ , or alter / , respectively — banking angles to 45 degrees where appropriate [27].

4 Network Evolution

On high-resolution small-scale displays such as paper, multiple dyad gestalts can be combined easily to enable visual exploration of patterns, trends, and outliers in complete longitudinal network data by simple means:

4.1 Matrix Representation

Their character-like representation enables a straightforward integration of dyad gestalts into what we refer to as a *gestaltmatrix*. Each matrix entry m_{ij} depicts a single dyad gestalt, i.e. the evolution of the complete dyadic time-series of ego i and alter j. Note that matrix entries m_{ij} and m_{ji} are redundant, since depicted gestalts are just mirrored at the vertical axes. The full matrix is displayed nevertheless to capitalize on the alignment across rows and columns. It facilitates, therefore, the detection of overall patterns, trends, and outliers from an ego point of view (the row) or from the view of all others (the column), respectively.

To the best of our knowledge, the most similar approach to representing change in social networks has been proposed in [47]. There, the authors indicate time through various subdivisions of matrix cells, mapping each data value to a single colored pixel; relation $\mapsto$ color. Such an atomistic treatment is highly efficient in terms of visualizing "the largest amount of data which is possible on current displays" [32]. In terms of gestalt laws, however, a holistic mapping – relation $\mapsto$ angle & length – seems to be more appropriate for indicating extent and balance (cf. Figure 5). Other related approaches involve animation and interaction, such as toggling the colors of matrix cells according to a queried time range [23].

4.2 Matrix Ordering

The crucial degree of freedom in a matrix-based network representation is the ordering of actors which determines the permutation of rows and columns. In social network analysis, orderings are often determined to highlight higher level organization in the matrix. The corresponding technique of blockmodeling (see, e.g., [14]) refers to substantively meaningful rearrangements that (visually) reveal regularities of the network structure within the matrix cells. For instance, one may be interested in cohesive groups and thus order the matrix so that locally dense groups form blocks along the diagonal. While this is not the only criterion, it is certainly among the one most commonly used. For most scenarios, however, finding an optimal permutation is $\mathcal{NP}$-hard. Various heuristics to calculate acceptable solutions for given criteria have been proposed. A comprehensive overview of blockmodeling techniques is provided in [14].

Since we aim for a single matrix-like diagram conveying the entire network evolution, we here define an aggregated actor similarity that, subsequently, can be used as the input to a blockmodeling approach. In other words, we describe a dyad by some similarity measure s that summarizes the time-series information $m_{ij} = s\big((d^1_{ij}, d^1_{ji}), \dots, (d^T_{ij}, d^T_{ji})\big)$.

The quality of an ordering obtained from such a similarity depends on its suitability for the analytic perspective. As a default that can be used in the absence of more specific requirements, we propose to sum the geometric means of dyadic ratings, $m_{ij} = \sum_{t=1}^{T} \sqrt{d^t_{ij} \cdot d^t_{ji}}$. This represents a reasonable general-purpose solution because of its tendency to group actors with consistently high and balanced weights. Part of the rationale is the correspondence of the geometric mean with the amount of ink and slope used in the seesaw design.

5 Case Study

We illustrate and further elaborate the proposed principles on one of the most well-known longitudinal data sets in social network analysis. Being the subject of many previous analyses, the data is ideally suited to demonstrate our method's capability to provide previously unexplored insights.

5.1 Data

The data comprises complete sociometric preference rankings – 'like best' to 'like least' – among 17 previously unacquainted male students at the University of Michigan in the mid 50s. The rankings were collected in 15 consecutive weeks and have been published in [40, 42]. Since all participants got free accommodation in fraternity housing, the data is commonly referred to as *Newcomb Fraternity Data*.

The development of friendships in the Newcomb Fraternity Data has been (re)analyzed numerous times, e.g., in [52, 35, 15, 38, 58]. Notably, it has also been used as an illustrating example in the introduction of stochastic actor-oriented models for dynamic network analysis [45].

Most descriptions of the Newcomb Fraternity Data have been based on summary measures, data aggregations and cross-sectional network views as in Figures 2 and 3. Beyond that, various animations between single states have been presented, e.g., in [35]. Even with the most sophisticated animations, though, attempting to describe and compare the evolution of multiple relationships or detecting interesting patterns – such as non-requited friendship nominations – inevitably results in cognitive overload.

The question is, thus, whether an analyst can gain further insights from the proposed static views on the overall evolution.

5.2 Exploring Individual Relations

Throughout the observation span, most mutual rankings in the Newcomb Fraternity Data are highly inconsistent, which indicates a low level of reciprocity or a high level of imbalance, respectively. This "large number of asymmetric ties suggests that we might gain some insight by using a layout method that accounts for this asymmetry" [35, page 1228]. Given highly asymmetric relational data, however, directed (node-)link representations become less comprehensible already in the static case [30].

In contrast, the proposed static gestalt-based representation provides a convenient and intuitive way to explore and communicate extent and balance in asymmetric relations. Because of their small, letter-like space requirements, the graphics can be used directly inside of (publishable) text. In this way, it is not necessary to go back and forth between an explanation and a diagram — facilitating comprehension, improving readability and conveying more information. For illustration compare the two paragraphs in Figure 6 with regard to the entropy of information.

For example, 7 12 are friends almost from the beginning. While 2 befriends 17 only late, 16 17 seems to stand no chance at that. The relationships of 7 14, 3 15, or 10 15 are rather more complicated. The latter is no surprise as 10 is unpopular with many.

"Nodes 10 and 15, for example, quickly emerge as nodes on the edge of the social structure. While they nominate each other symmetrically early in the observation period, they lose interest in each other by the end. Neither node receives top-five nominations from any other node in the network. Their nominations to others seem to dance around the graph, never resting for long on a single person." [35, p. 1228]

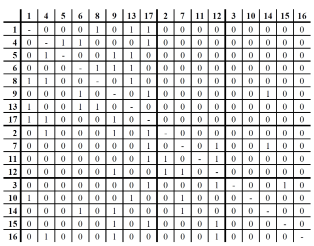
Fig. 6. Explanation with datawords: more information with fewer words.

5.3 Exploring Groups of Relations

Besides the persistent asymmetry in friendship rankings and large heterogeneity in the popularity of actors (addressed below), the evolution of group structures within the Newcomb Fraternity Data has been studied extensively.

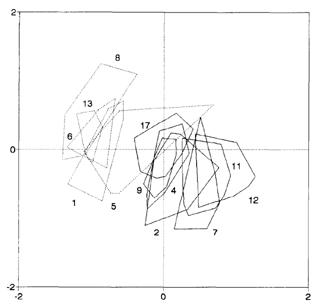
Evaluating blockmodeling algorithms according to the density of inter- and intra-block top ratings, e.g., White et al. state that “by at least the fifth week not only the final blocks but also the final blockmodel have emerged with remarkable clarity”[58, page 764]. Their findings basically confirm the actor groups originally proposed by Nordlie [42] based on rank correlations. Nakao and Romney [38] provide further evidence for structural convergence in the Newcomb Fraternity Data by relating the number of concordant (c) and discordant (d) rankings with Goodman-Kruskal’s Gamma coefficient $(c-d)/(c+d) \in [-1,1]$.

Findings on group structures in the Newcomb Fraternity Datahave been presented either textually (13 9 17 1 8 6 4) (7 11 12 2) (14 3 10 16 5 15) [58, page 764] or by depicting a permuted adjacency matrix; compare Figure 7(a). An exception can be found in [38], where the authors use multidimensional scaling to place actors in the plane according to the similarities of their rankings at a given point in time. Then, Procrustes analysis is used to align each actor’s positions over time, as indicated by the convex hulls presented in Figure 7(b).



	1	4	5	6	8	9	13	17	2	7	11	12	3	10	14	15	16
1	-	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0
4	0	-	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0
5	0	1	-	0	0	1	1	0	0	0	0	0	0	0	0	0	0
6	0	0	0	-	1	1	1	0	0	0	0	0	0	0	0	0	0
8	1	1	0	0	-	0	1	0	0	0	0	0	0	0	0	0	0
9	0	0	0	1	0	-	0	1	0	0	0	0	0	0	1	0	0
13	1	0	0	1	1	0	-	0	0	0	0	0	0	0	0	0	0
17	1	1	0	0	0	1	0	-	0	0	0	0	0	0	0	0	0
2	0	1	0	0	0	1	0	1	-	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	1	0	-	0	1	0	0	1	0	0
11	0	0	0	0	0	0	0	1	1	0	-	1	0	0	0	0	0
12	0	0	0	0	0	1	0	0	1	1	0	-	0	0	0	0	0
3	0	0	0	0	0	0	0	1	0	0	0	1	-	0	0	1	0
10	1	0	0	0	0	0	1	0	0	1	0	0	0	-	0	0	0
14	0	0	0	1	0	1	0	0	0	1	0	0	0	0	-	0	0
15	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	-	0
16	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	-

(a) wave 14 tabular [52, page 154]



(b) MDS + Procrustes [38, page 122]

Fig. 7. Previous descriptions of groups in Newcomb Fraternity Data.

Any grouping can be visualized compactly and evaluated in a single *gestaltmatrix* without data aggregation. For illustration we here use an ordering based on the number of inversions in consecutive rankings  , which assumes local minima in week 8 and week 12. Thus, the ratings in week 7 and week 11 are rather stable and, therefore, reasonable proxies for static representations. We choose week 11 and sort actors according to their values in the Fiedler vector of the Laplacian matrix obtained from (m_{ij}), where $m_{ij} = \sqrt{d_{ij}^{11} \cdot d_{ji}^{11}}$ if both ratings are top 3-ratings and $m_{ij} = 0$ otherwise. This choice of similarity measure is motivated by the common use of thresholding in published reanalyses of the Newcomb Fraternity Data; compare Figures 2 and 3. The resulting gestaltmatrix is presented in Figure 8. A block structure due to higher internal rankings is clearly visible.

As suggested in Section 3.3 we augment the plain color design. First, inspired by the gestalt principle of *figure and ground*, we combine two different hues: using black for highlighting top 3 values of the extent and gray for fading the others, we intend to assess the result of the commonly used thresholding before the background of the full data set. Second, motivated by the gestalt principle of *similarity*, we distinguish different types of actors by coloring them according to their standardized popularity over time. Popular actors are colored blue, whereas unpopular actors are colored red. More specifically, we evaluated the deviation π from a null model, assuming that at each time step, each actor obtains each possible ranking at a time, i.e. $\mu = \sum_{j \neq i} d_{ji}^t / k = k/2$, with $k = \#\{j \neq i\}$ and assuming rankings from k to 1. This gives

$$\pi(i) = \frac{1}{T} \cdot \sum_{t=1}^{T} \sum_{j \neq i} \frac{d_{ji}^t / k - \mu}{\sigma}$$

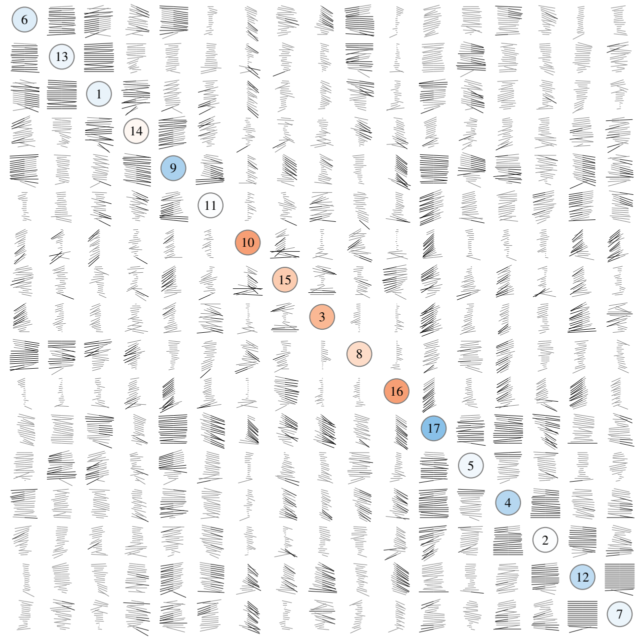


Fig. 8. Gestaltmatrix of Newcomb Fraternity Data, showing the evolution using all 4080 data points (rankings) from 15 waves. Labels on the diagonal are numbered according to Nordlie [42] and colored according to the standardized overall popularity of the corresponding actor. The sorting of actors is according to a spectral approach described in the main text.

where σ refers to the standard deviation of a single ranking d_{ji}^t from μ and T is the number of panels.

A variant gestaltmatrix with identical coloring is shown in Figure 10. In this matrix we reproduce an ordering of the actors proposed in [38]. With the gestaltmatrix view on the complete evolution of relations, however, additional insights can be obtained from the more refined perspective. Consider the column of the “scapegoat in this group (man 10), who received one of the bottom three choices of each of the other persons” [58, p. 759]. It is clearly visible which others did overrate the ‘scapegoat’ in the beginning. Interestingly, these include the overall most popular actors 17, 4 and 9. Also, a strong desire of the most unpopular actors to be friends especially with the popular actors is revealed. To the best of our knowledge, these findings on the Newcomb Fraternity Data have never been published before, despite long-standing investigations spanning various disciplines.

As a concluding example, we stress once more the potential to complement and critically inform existing analyses, based on animation or interaction. Figure 9 demonstrates the consequences.

“For example, one can see that nodes 1, 6, 8, and 13 remain strongly connected to each other throughout the observation period, occupying a small cluster at the right of the graph. Nodes 7, 12, and 4 form a cluster early in the groups history, but node 4 then breaks with this group at about week 8, instead nominating nodes 17 and 2.” [35, p. 1228]

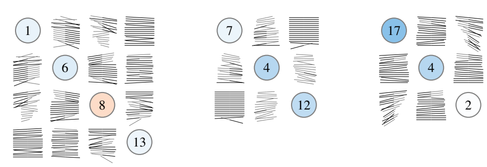


Fig. 9. Illustration with the actual data casts doubt on statements from the literature.

Fig. 10. Gestaltmatrix of Newcomb Fraternity Data, showing the complete evolution of 4080 rankings collected in 15 waves. Labels on the diagonal are numbered according to Nordlie [42] and colored according to standardized overall popularity of corresponding actor. Arrangement of actors according to the sorting proposed in [38].

6 Expert Feedback

The representation proposed here is intended for domain expert use and scientific communication. While it is currently being used in several cooperations, we may not show proprietary data in this paper. This is one reason why, in the case study of Section 5, we resorted to a well known and fairly typical data set which, in addition, has been the subject of many previous analyses.

We will ultimately be interested in a quantitative comparison assessing the relative efficiency of our design in communicating extent and balance in asymmetric relations. In the absence of sufficiently similar previous approaches, however, the number of dimensions and confounding factors appears to be too large for a formal, well-focused user study to be included in this paper [11]. Instead we rely on external expert reviews [51] as an initial sanity check.

Because of its intuitiveness and some other obvious benefits (static representation, generality, printability, publishability) we did assume that social scientists are willing to spend the five minutes it takes to understand the design and learn how to read it. The method was therefore presented at INSNA Sunbelt XXXI Social Networks Conference,[1] the main venue for networks in the social sciences (no published proceedings). Interest was surprisingly high and feedback overwhelmingly positive; even to the point that the idea was taken up and applied in another presentation at the same conference within two days.

Fourteen senior domain experts and a number of other delegates provided personal feedback in private and informal face-to-face discussions, with all but three approaching us before we could ask them to. They mentioned in particular the simplicity, intuitiveness, and aesthetic appeal of the design, and the ease by which tendencies and outliers can be detected both on the actor and dyad level. They also mentioned the compact and simultaneous representation of the entire data set, allowing back and forth comparison of matrix rows and cells. Generally, they were aware of the fact that distances between vertices are difficult to determine in matrix representation, but needed help in understanding the consequences of row and column reordering and the reliability of conclusions drawn from a given ordering. The latter was somewhat surprising because this aspect of our design is shared with any matrix representation of networks. Under the name sociomatrix the latter are known well in this domain [36, 19].

Feedback was particularly vocal from the modeling community, where the interest is in identifying factors that govern the evolution of dyads [46]. Detecting evolving patterns and, in particular, exceptional actors, is of great importance in these approaches because current models rely on fairly strong homogeneity assumptions. The gestalt-based design was found to serve this purpose better than anything our respondents had known before.

These impressions were confirmed during a dedicated modeling workshop,[2] where we worked with two groups to represent their data appropriately in our design. Encouraged by the reactions so far, we plan to make the approach available in more generally accessible form, possibly in our tool visone.[3]

7 Discussion

We proposed a novel methodology for static visualization of longitudinal asymmetric network data, which we refer to as *gestaltmatrix*. It is based on word-sized representations of dyadic evolution, which in turn represent an application of gestaltlines, i.e. multivariate sparklines capitalizing on gestalt principles.

While, currently, scepticism with regard to glyph-based approaches to representing multidimensional data appears to dominate, our work is in line with other recent work such as [41], in which it is suggested that additional perceptually-effective forms of compact multidimensional representation may await discovery and characterization.

Since our approach provides novel means to explore and communicate the extent and balance of values in dyads, it can also be used to complement existing techniques that require animation or interaction. Next to some obvious benefits such as effective communication of findings and suitability for publication on paper we also demonstrated that gestalt-based network analysis bears the potential to yield additional insight even into data that was previously studied extensively. Additionally, we revealed misinterpretations that we suspect resulted from more aggregate data views.

There are many ways in which additional information about, e.g., volatility or other attributes can be integrated into a gestaltmatrix design. We have only shown one example by indicating the average popularity of each actor on the diagonal. The main goal was to argue that a detailed design can lead to interpretable forms on the level of the dyad (matrix cell), the actor (row), and the network (matrix). These forms are likely to ease the discovery of trends, change events, and outliers.

There are two major limitations for the scope of our method. The first one is shared with all matrix representations of networks, namely that paths are difficult to discover and follow. The other one is a consequence of our attempt to bring out a joint appearance of the data in a dyad; as a consequence, individual time slices are difficult to extract. Possible remedies for these two problems may lie in a combination of the seesaw metaphor with node-link diagrams and by using gestaltlines highlighting the current point in time as edge labels. These, however, require more careful research. In ongoing collaborations with social scientists we are dealing with specialized data sets that contain in excess of 100 actors. This can be considered an upper limit for the large majority of empirical studies also from the modeling point of view. For larger networks, however, larger print and, possibly, a hierarchical design may become necessary.

Qualitative evidence on the acceptance and intuitive understanding of the proposed principles was provided based on informal feedback from domain experts. Further research shall include a quantitative assessment of the impact that gestaltmatrices have on the understanding of asymmetric relations in longitudinal social networks.

[1] 8-13 February 2011, St. Pete Beach, FL, USA; approx. 550 delegates.

[2] 1st Advanced Siena Users Meeting (AdSUM 2011), April 7+8, Konstanz; 41 delegates.

[3] `www.visone.info`

Acknowledgments

This work was performed in part within the argo Transfer Center for Network Research (University of Konstanz, FP 665/10). We thank the anonymous reviewers for their helpful and constructive comments.

References

[1] G. D. Battista, P. Eades, R. Tamassia, and I. G. Tollis. *Graph Drawing: Algorithms for the Visualization of Graphs*. Prentice Hall, 1999.

[2] S. P. Borgatti and D. S. Halgin. Network theorizing. *Organization Science*, 2011. To appear.

[3] S. P. Borgatti, A. Mehra, D. J. Brass, and G. Labianca. Network analysis in the social sciences. *Science*, 323(5916):892–895, 2009.

[4] U. Brandes, N. Indlekofer, and M. Mader. Visualization methods for longitudinal social networks and stochastic actor-oriented modeling. Submitted to Social Networks, 2011.

[5] U. Brandes, P. Kenis, and J. Raab. Explanation through network visualization. *Methodology*, 2(1):16–23, 2006. Spanish translation in *REDES* 9(6), 2005.

[6] U. Brandes, P. Kenis, and D. Wagner. Communicating centrality in policy network drawings. *IEEE Transactions on Visualization and Computer Graphics*, 9(2):241–253, 2003.

[7] U. Brandes, J. Raab, and D. Wagner. Exploratory network visualization: Simultaneous display of actor status and connections. *Journal of Social Structure*, 2(4), 2001.

[8] U. Brandes and D. Wagner. A Bayesian paradigm for dynamic graph layout. In G. Di Battista, editor, *Proceedings of the 5th International Symposium on Graph Drawing (GD'97)*, volume 1353 of *Lecture Notes in Computer Science*, pages 236–247. Springer-Verlag, 1997.

[9] J. M. Chambers, W. S. Cleveland, B. Kleiner, and P. A. Tukey. *Graphical Methods for Data Analysis*. Chapman and Hall, 1983.

[10] D. Chang, L. Dooley, and J. E. Tuovinen. Gestalt theory in visual screen design: a new look at an old subject. In *CRPIT '02: Proceedings of the 7th world conference on computers in education*, 2002.

[11] R. K. Christopher, C. G. Healey, V. Interrante, D. H. Laidlaw, and C. Ware. Thoughts on user studies: Why, how, and when. *IEEE Computer Graphics and Applications*, 23:20–25, 2003.

[12] C. D. Correa and K.-L. Ma. Visualizing social networks. In C. C. Aggarwal, editor, *Social Network Data Mining*, pages 307–326. Kluwer Academic Publishers, 2011.

[13] J. Díaz, J. Petit, and M. Serna. A survey of graph layout problems. *ACM Computing Surveys*, 34(3):313–356, 2002.

[14] P. Doreian, V. Batagelj, and A. Ferligoj. *Generalized Blockmodeling*. Cambridge University Press, 2005.

[15] P. Doreian, R. Kapuscinski, D. Krackhardt, and J. Szczypula. A brief history of balance through time. *Journal of Mathematical Sociology*, 21:113–131, 1996.

[16] N. Elmqvist, T.-N. Do, H. Goodell, N. Henry, and J.-D. Fekete. ZAME: Interactive large-scale graph visualization. In *Proc. IEEE Pacific Visualization Symposium (PacificVis 2008)*, pages 215–222, 2008.

[17] S. Few. *Information Dashboard Design : The Effective Visual Communication of Data*. O'Reilly, 2006.

[18] K. Flieder and F. Mödritscher. Foundations of a pattern language based on gestalt principles. In *CHI '06 extended abstracts on Human factors in computing systems*. ACM, 2006.

[19] E. Forsyth and L. Katz. A matrix approach to the analysis of sociometric data: preliminary report. *Sociometry*, 9:340–347, 1946.

[20] L. C. Freeman. Visualizing social networks. *Journal of Social Structure*, 1(1), 2000.

[21] Y. Frishman and A. Tal. Online dynamic graph drawing. *IEEE Transactions on Visualization and Computer Graphics*, 14(4):727–740, 2008.

[22] M. Ghoniem, J.-D. Fekete, and P. Castagliola. On the readability of graphs using node-link and matrix-based representations: a controlled experiment and statistical analysis. *Information Visualization*, 4:114–135, 2005.

[23] M. Ghoniem, N. Jussien, and J.-D. Fekete. Visexp: visualizing constraint solver dynamics using explanations. In *FLAIRS'04: 17th intl Florida Artificial Intelligence Research Society conf.* AAAI press, 2004.

[24] B. D. Greenhill, M. J. Ward, and A. E. Sacks. The 'separation plot': A new visual method for evaluating the predictive power of logit/probit models. APSA 2009 Toronto Meeting Paper. Available at SSRN: http://ssrn.com/abstract=1450589, 2009.

[25] M. S. Handcock and K. J. Gile. Modeling social networks from sampled data. *Annals of Applied Statistics*, 4(1):5–25, 2010.

[26] M. Harrower and C. A. Brewer. ColorBrewer.org: an online tool for selecting colour schemes for maps. *The Cartographic Journal*, 40(1):27–30, 2003.

[27] J. Heer and M. Agrawala. Multi-scale banking to 45 degrees. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):701–708, 2006.

[28] J. Heer, N. Kong, and M. Agrawala. Sizing the horizon: The effects of chart size and layering on the graphical perception of time series visualizations. In *In Proc. ACM Human Factors in Computing Systems (CHI)*, pages 1303–1312, 2009.

[29] N. Henry, J.-D. Fekete, and M. J. McGuffin. NodeTrix: A hybrid visualization of social networks. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1302–1309, 2007.

[30] D. Holten and J. J. van Wijk. A user study on visualizing directed edges in graphs. In *Proceedings of the 27th international conference on Human factors in computing systems*, CHI '09, pages 2299–2308, 2009.

[31] M. Kaufmann and D. Wagner, editors. *Drawing Graphs: Methods and Models*. Springer, 2001.

[32] D. A. Keim. Designing pixel-oriented visualization techniques: Theory and applications. *IEEE Transactions on Visualization and Computer Graphics*, 6(1):59–78, 2000.

[33] G. Kossinets. Effects of missing data in social networks. *Social Networks*, 28(3):247–268, 2006.

[34] P. V. Marsden. Network data and measurement. *Annual Review of Sociology*, 16:435–463, 1990.

[35] J. Moody, D. A. McFarland, and S. Bender-DeMoll. Dynamic network visualization. *American Journal of Sociology*, 110(4):1206–1241, 2005.

[36] J. L. Moreno. *Who Shall Survive? Foundations of Sociometry, Group Psychotherapy, and Sociodrama*. Beacon House, 1953. Originally published in 1934.

[37] M. Morris. Overview of network survey designs. In M. Morris, editor, *Network Epidemiology: A handbook for survey design and data collection*, pages 8–24. Oxford University Press, 1st edition, 2004.

[38] K. Nakao and A. K. Romney. Longitudinal approach to subgroup formation: re-analysis of newcomb's fraternity data. *Social Networks*, 15(2):109–131, 1993.

[39] K. V. Nesbitt and C. Friedrich. Applying gestalt principles to animated visualizations of network data. In *Proc. Information Visualisation*, pages 737–743, 2002.

[40] T. M. Newcomb. *The Acquaintance Process*. New York: Holt, Rinehart & Winston, 1961.

[41] C. B. Nielsen, S. D. Jackman, I. Birol, and S. J. M. Jones. Abyss-explorer: Visualizing genome sequence assemblies. *IEEE Transactions on Visualization and Computer Graphics*, 15:881–888, 2009.

[42] P. Nordlie. A longitudinal study of interpersonal attraction in a natural group setting. Unpublished Ph.D. dissertation, University of Michigan, 1958.

[43] A. Perer and B. Shneiderman. Balancing systematic and flexible exploration of social networks. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):693–700, 2006.

[44] R. Sedgewick. *Algorithms in C*. Addison-Wesley, 1998.

[45] T. Snijders. Stochastic actor-oriented models for network change. *Journal of Mathematical Sociology*, 21:149–172, 1996.

[46] T. A. Snijders, G. G. van de Bunt, and C. E. Steglich. Introduction to stochastic actor-based models for network dynamics. *Social Networks*, 32(1):44–60, 2010.

[47] K. Stein, R. Wegener, and C. Schlieder. Pixel-oriented visualization of change in social networks. In *Advances in Social Networks Analysis and Mining*, pages 233–240, 2010.

[48] R. J. Sternberg. *Cognitive Psychology*. Wadsworth Publishing, 5th edition, 2008.

[49] S. Thakur and T.-M. Rhyne. Data vases: 2d and 3d plots for visualizing multiple time series. In *Proc. of the 5th International Symposium on Advances in Visual Computing: Part II*, pages 929–938, 2009.

[50] A. Tikhonova, C. D. Correa, and K.-L. Ma. Visualization by proxy: A novel framework for deferred interaction with volume data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1551–1559, 2010.

[51] M. Tory and T. Moller. Evaluating visualizations: Do expert reviews work? *IEEE Computer Graphics and Applications*, 25:8–11, 2005.

[52] M. Trappmann, H. J. Hummell, and W. Sodeur. *Strukturanalyse sozialer Netzwerke*. VS Verlag, 2nd edition, 2011.

[53] E. R. Tufte. *Envisioning Information*. Graphics Press, 1990.

[54] E. R. Tufte. *Beautiful Evidence*. Graphics Press, 2006.

[55] C. Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann, 2nd edition, 2004.

[56] M. Wattenberg. Visual exploration of multivariate graphs. In *Proc. ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2006)*, pages 811–819, 2006.

[57] M. Wertheimer. Untersuchungen zur Lehre von der Gestalt II. *Psychologische Forschung*, 4:301–350, 1923.

[58] H. C. White, S. A. Boorman, and R. L. Breiger. Social structure from multiple networks. I. Blockmodels of roles and positions. *The American Journal of Sociology*, 81(4):730–780, 1976.

[59] J.-S. Yi, N. Elmqvist, and S. Lee. TimeMatrix: Visualizing temporal social networks using interactive matrix-based visualizations. *International Journal of Human-Computer Interaction*, 26(11 & 12):1031–1051, 2010.