19th International Congress on Modelling and Simulation, Perth, Australia, 12–16 December 2011
http://mssanz.org.au/modsim2011

Temporal Social Network Analysis of Discourse

A.H. Dekker

Defence Science and Technology Organisation Australia
Email: dekker@acm.org

Abstract: In this paper, we explore a number of techniques for two-dimensional visualisation of temporal social network data, with the goal of providing feedback on the group dynamics of government planning workshops. These techniques include the use of position and colour for displaying temporal information, as in colour-coded bar charts or sequence diagrams. While it is possible to display temporal information solely using colour, experiments with an Internet discussion group showed that the most useful techniques are plots with an explicit horizontal time axis, and a vertical indication either of the amount of communication by actors (weighted degree), or of the importance of actors in the patterns of interaction (centrality).

We have applied these visualisation techniques to interaction data from a workshop case study (Fleming, 2008). Doing so identified *scaled centrality plots*, such as that in Figure 1, as particularly useful for providing insights into the changing patterns of interaction between workshop participants. These plots display actor centrality calculated over a sliding window, scaled by the average centrality in that window, and are useful for providing feedback to workshop facilitators on the overall "flow" of their workshop over time.

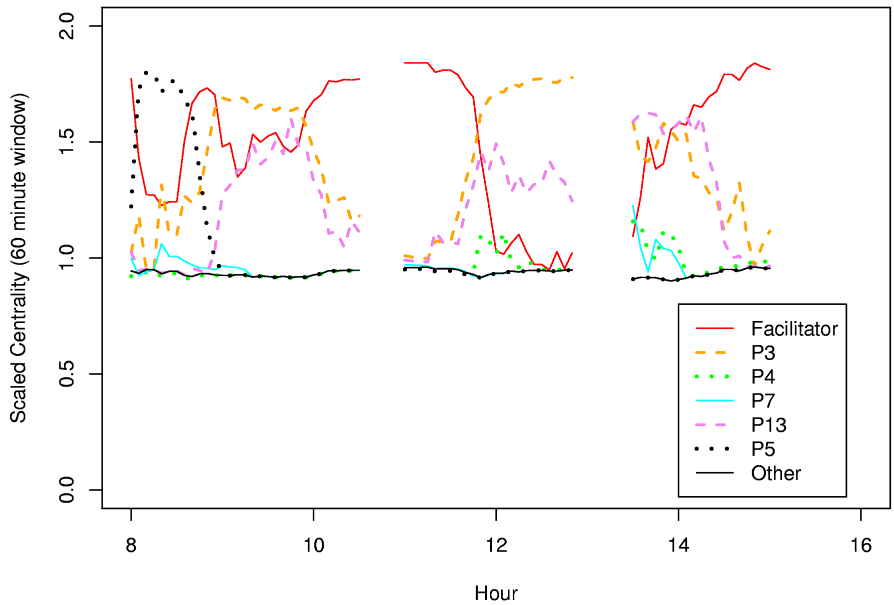

Figure 1. Scaled centrality plot for our workshop case study. The nine sub-phases of the workshop (discussed in Section 4) stand out clearly, making such plots helpful in visualising workshop dynamics. In particular, there is a very brief initial dominance by the facilitator (setting the scene), followed by a period dominated by participant P5, then by renewed facilitator control, then by a discussion dominated by P3 and P13, then by five other sub-phases. The explicit time axis assists in communicating the information in this diagram to workshop facilitators.

Keywords: *social network analysis, visualisation, workshops, discourse analysis*

1. INTRODUCTION

Workshops form an important technique for helping teams of people in government and business come to grips with difficult problems. Support can be provided to such workshops by observing and analysing the interactions that take place and the changes in interaction patterns over time. Analysis of these interaction patterns provides a road to understanding the group dynamics of the workshop, which in turn can provide valuable feedback to workshop facilitators.

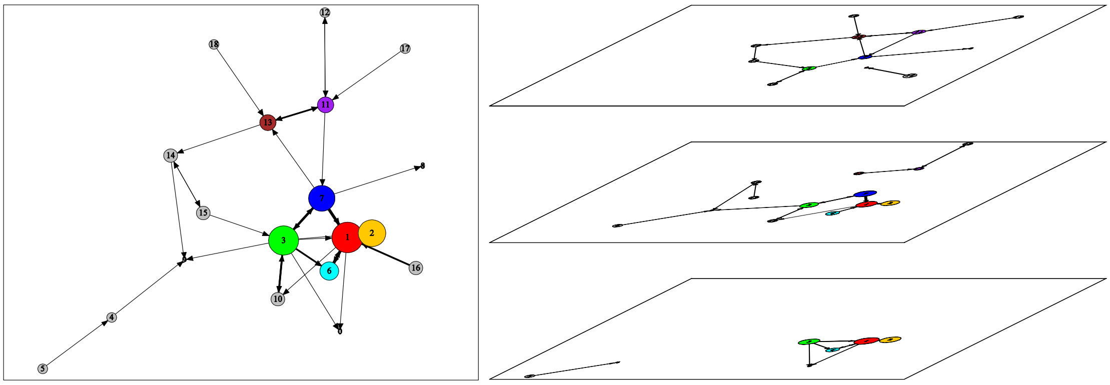

Figure 2. Social network for the Internet discussion group `sci.math`, produced by spring-embedding, with 7 key participants colour-coded (left). Node area and arrow thickness indicate number of interactions during the sample period. On the right are three temporal "slices" of the social network, with time running upwards.

Social Network Analysis is a technique which sheds light on the analysis of workshop dynamics and other organisational processes. Traditionally, social networks are essentially 2-dimensional objects – represented as two-dimensional matrices or as two-dimensional visualisations, as in the left of Figure 2 (Wasserman and Faust, 1994; Freeman, 2000; Dekker, 2001). Such images have several advantages for describing social networks – they can be printed on paper or displayed on computer screens, and can act as a focal point for discussion of the interactions in the network.

Adding a temporal element (e.g. Alice communicates to Bob at 10:00 AM, and to Charles at 10:05) makes social networks essentially 3-dimensional. By dividing time into a series of "slices" (Bender-deMoll and McFarland, 2006) a temporal social network can be viewed as a series of static social networks, as in the right of Figure 2. Such a temporal series can be converted into an animation or video (Bender-deMoll and McFarland, 2006), and viewing such an animation can give a good picture of the dynamic evolution of the social network (Fleming, 2007). However, such animations cannot be printed, and are not always suitable as a focal point for discussion. They can be displayed on a computer screen, but highlighting a key feature requires fast-forwarding and then pausing (which removes the temporal element).

How can we visualise the temporal evolution of a social network while retaining the advantages of a 2-dimensional representation? In this paper, we provide a taxonomy of techniques for doing so (Section 2), illustrated via a case study involving comments in the Internet discussion group `sci.math`. Sections 3 and 4 follow with an analysis of observational data on workshop dynamics, collected by Fleming (2008). Our focus in this paper is primarily on descriptive visualisations that can be combined with other observational data to give feedback to workshop facilitators. We do not address inferential approaches to temporal network data, such as that of Snijders (2005).

2. TEMPORAL VISUALISATION OF SOCIAL NETWORKS

Our taxonomy of visualisation techniques for temporal social networks is illustrated with a dataset of interactions collected from the Internet discussion group `sci.math`, time-stamped between 15:59 on 13 March 2008 (Australian Eastern Daylight Time) to 04:11 on 14 March 2008 (which we will treat for analysis as the nonstandard time 28:11), illustrated in Figure 2. The dataset includes only postings which explicitly comment on postings by another author, and which therefore can be viewed as comments to the author in question. It is therefore a directed temporal social network.

There are two fundamental approaches we can take to producing 2-dimensional temporal diagrams of such discussions. The first is to take time as one axis of the diagram, in which case the other must display some 1-dimensional characterisation of the network at a particular point in time. That is, the other axis must show numerical properties of the actors (e.g. *degree* and *centrality*), of the links, or of the network as a whole. The other approach is a 2-dimensional representation with no time axis, where temporal information is somehow encoded in specific attributes of the diagram – either in attributes of the nodes or of the links. Possible attributes include *position* and *colour*, and we will consider one example of each option.

2.1. Time Axis with Degree

The first form of visualisation we consider is a 2-dimensional diagram where one axis represents time, and the other represents (weighted) actor out-degrees. The out-degree of an actor for a specific time interval is the number of interactions (i.e. comments) by that actor during the interval, which can be conveniently displayed as a histogram or bar chart, as in Figure 3. In this chart, one-hour time intervals are used, and interactions by key participants (those responsible for at least three interactions) are colour-coded, while those by minor participants are shown in grey.

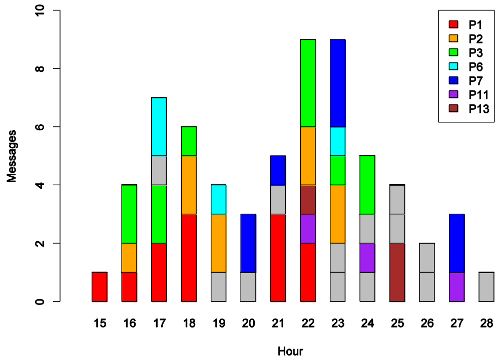

Figure 3. Bar chart of interactions for `sci.math` case study, colour-coded as in Figure 2. Each bar indicates the number of messages for a one-hour period starting at the indicated time (with 24 as midnight and 25 as 01:00 the next morning).

The bar chart in Figure 3 gives a good idea of the ebbs and flows of communication in this discussion forum during the sample period. For example, Person 1 (P1, in red) sent 12 messages (19% of the total), finishing at 22:56, in line with the end of the workday in Europe, where P1 was resident.

A bar chart of this kind is familiar to most people, and is very easy to explain and understand. However, while it shows how much each participant contributed, and when, it gives no indication of network structure. It does not, for example, show that 75% of messages from P1 were in response to one other participant (P2).

2.2. Time Axis with Arrow Counts

It is possible to address some of the limitations of Figure 3 by colour-coding the bar chart by arrows (dyads) rather than by actors. Doing so reveals that 46% of the interactions occurred on only five arrows: P1 to P2; P2, P6, and P7 to P1; and P3 to P7. However, this approach becomes unwieldy for networks with many links.

2.3. Sequence Diagrams

Another alternative to Figure 3 is to place the interactions themselves along the time axis. In the software industry, this is known as a *sequence diagram*, and is used to show interactions between software objects (Fowler and Scott, 1997). In such a diagram, interactions are shown as vertical arrows between nodes representing actors. For small networks, a sequence diagram gives considerable information about the temporal evolution of the network, but with many interactions, such diagrams can become cluttered.

2.4. Time Axis with Centrality

Centrality is a numeric characteristic of actors that, unlike degree, takes network structure into account (Wasserman and Faust, 1994). Because centrality scores are only meaningful if structure is indeed present, they must be calculated for a specific time period. It is convenient to perform the calculation within a sliding time window, as in Howison *et al.* (2006). Figure 4 shows centrality scores for the seven key actors from Figure 2, calculated over a two-hour sliding time window. The position along the time axis marks the midpoint of the window over which centrality was calculated.

The network of interactions with a given time window will in general not be connected. It is therefore important to use a definition of centrality that permits disconnected networks, in which some distances $d(x,y)$ are infinite. We use the *valued centrality* measure of Dekker (2005), which correlates well with the traditional closeness centrality when networks are connected, while allowing disconnected networks (Dekker, 2008).

The valued centrality of an actor x is defined as the average of reciprocal distances $d(x,y)$ to all other nodes (if there is no link, $d(x,y) = \infty$, where $1/\infty = 0$):

$$C_V(x) = \frac{1}{n-1}\left(\sum_{y \neq x} \frac{1}{d(x,y)}\right) = \underset{y \neq x}{AVG}\left(\frac{1}{d(x,y)}\right)$$

The centrality plot in Figure 4 clearly shows the three main phases of discussion mentioned above, and highlighted in the three network diagrams at the top of the chart. These network diagrams, also shown in Figure 2, indicate interactions for the periods 15:59 to 19:30, 19:30 to 24:00, and 24:00 to 04:11 (28:11).

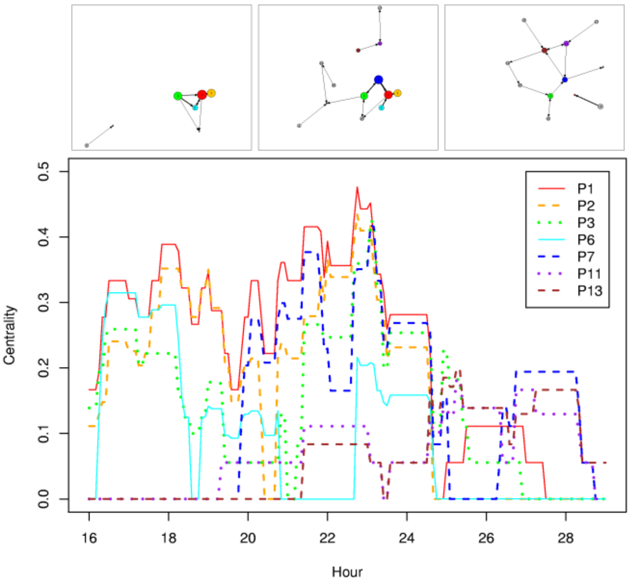

Figure 4. Centrality plot for key participants in the `sci.math` case study, showing three phases of discussion, corresponding to the network diagrams at the top of the chart, which are also the network slices on the right in Figure 2.

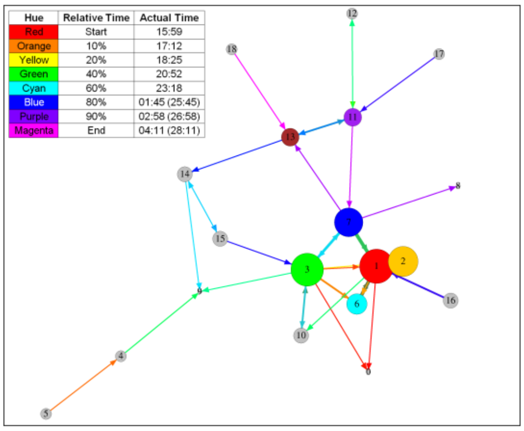

Figure 5. Social network for `sci.math`, using colour to indicate time of activity on arrows. Hue indicates average time (the legend shows eight standard colours), and saturation indicates time spread. Node position is as for Figure 2.

Sliding-window centrality plots like Figure 4 are particularly powerful in identifying key participants at various points in time, and in identifying discussion phases which can be explored in more detail using network diagrams. For this example, the phases of discussion result from a trivial factor: the time zones in which participants reside. However, in our workshop study (see Section 4), the pattern of discussion phases is more interesting.

An alternate use of centrality is the “temporal social surface” of Gloor (2005), which is a 3-dimensional plot obtained by giving, at each point in time, an ordered list of node centralities. These plots clearly illustrate changes in the number of highly central individuals. However, they do not indicate who these individuals are, and so are inadequate for our purposes.

2.5. Temporal Colouring of Arrows

As indicated above, it is possible to encode temporal information in a network diagram, without including an explicit time axis. One technique for doing so is to apply colour to the arrows in order to indicate the time at which activity took place, and this is done in Figure 5, where the hue of the arrows indicates the average time of activity, in a spectrum ranging from red to magenta. For example, the arrow from P1 to P0 (at 15:59) is red, the arrow from P1 to P6 (at 17:33) is very close to standard orange, and the arrow from P13 to P14 (at 01:32) to standard blue. The spread of times around the average value is indicated by colour saturation, with wider spreads being greyer in colour.

However, although this diagram does indeed display timing information, it harder to interpret than an explicit time axis, and we have found this method less useful in general than the others.

2.6. Temporal Correlation of Actors

The final approach we consider is the encoding of temporal information by position, other than using an explicit time axis. Specifically, we have used multi-dimensional scaling (Brandes, 2001) to display temporal similarity. The temporal similarities of actors are calculated by associating each message with a triangle of activation, starting an hour before message transmission, peaking at the transmission time, and dropping to zero an hour after transmission. Each actor’s activation history is the sum of the triangles for all messages sent, and the temporal activity similarity of two actors is the squared correlation of their activity histories.

This process gives a network diagram which groups together actors who are active at similar times. However, it gives no clear indication of the temporal sequence, and we have found this rather abstract representation of temporal evolution difficult to explain to non-specialists. It is therefore of limited utility for discourse analysis, except where there is clear clustering of actors. In general, time seems to be best represented using an explicit time axis, and we have used explicit time axes in analysing our workshop data.

3. WORKSHOP DATA COLLECTION

Fleming (2008) has collected discourse data from a number of government planning workshops, and data from one such workshop is analysed in Section 4. The data was collected by observation of participant interactions, using a software tool with a graphical user interface to record which person is speaking, and whether they were addressing another individual or the group as a whole. Figure 6 shows a simplified version of the tool used. The tool also permits a rough categorisation of utterances made. The interactions are recorded by clicking on buttons representing participants and topics, and so only a few mouse-clicks are required to record each interaction. Time stamps are provided automatically by the collection tool. In the event of data entry problems, the data can be corrected after the fact using video recordings of the workshop.

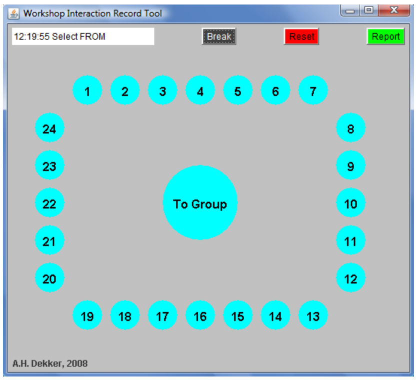

Figure 6. Simplified version of the data collection tool.

In this paper, we concentrate on network-based analysis of the workshop interactions, using the techniques of Section 2, and therefore do not consider the categorisation of interaction type. The specific workshop discussed in Section 4 was one of a sequence of workshops, in which all participants had over a decade's experience on the topic being discussed – although participants were at different levels of seniority within their organisation. Most participants knew each other prior to the workshop, and also knew the facilitator, who was a trusted senior member of the organisation.

4. THE MAJOR WORKSHOP CASE STUDY

We have used the techniques described in Section 2 as a way of visualising the temporal dynamics of the workshop described above. During the workshop, there were 997 interactions, ranging from brief interjections to speeches several minutes long. The mean duration of interactions was 21 seconds.

In a previous analysis, presented in Fleming (2007), we have noted that the number of interactions by workshop participants tends to vary considerably, following a power law distribution (Watts, 2003). This does not necessarily reflect lack of interest or expertise – certainly it did not for the workshop analysed here. Indeed, questions by the facilitator indicated that all participants remained engaged during the workshop. Figure 7 shows that, for this workshop, participation followed a power law with a truncated tail. This was encouraging: it indicated that the number of low-activity participants was less than in some other workshops. Excluding the workshop facilitator, 77% of participants had more than 5 interactions.

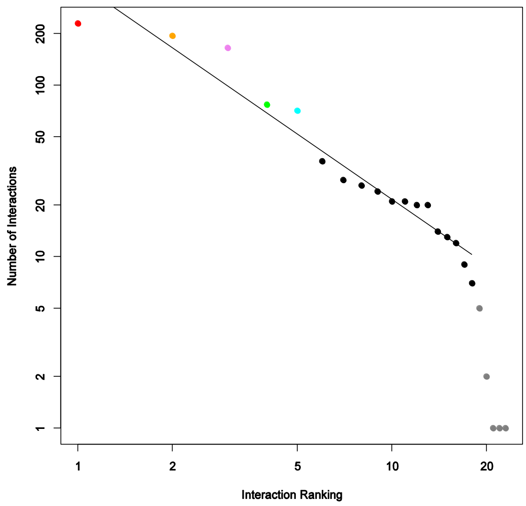

Figure 7. Interactions by these workshop participants followed a power law with a truncated tail.

The approximate power-law distribution of interactions means that it is generally useful to colour-code the most active participants. In this case, the facilitator (red) and four key participants (orange, violet, green, and cyan) were together responsible for 74% of all interactions. There is also a practical lesson here for workshop facilitators: the need to work constantly to encourage a more balanced pattern of participation, which corresponds to making the trend line in Figure 7 less steep.

For this workshop, the variation in the amount of participation was partly the result of differences in seniority level. On average (by geometric mean), senior staff interacted 4.7 times as much as junior staff, and this was significant at the 3% level (by analysis of variance on log-transformed data, F = 6.05). Such differences in the amount of participation represent a challenge for workshop facilitators.

Of the interactions during the workshop, most involved comments or statements to the group as a whole, and only 10% were directed at specific individuals. The network diagrams on top of Figure 8 show these directed interactions. The network is dominated by the five colour-coded participants, but on the top left there is also a visible "star" pattern of interaction involving Person 5 (P5), a senior stakeholder who helped to set the scene for the early part of the workshop. The bar chart in Figure 8 summarises all the interactions during the workshop. Interactions by the five key participants are colour-coded, but only interactions to the group as a whole. The dark blue portions at the top of each bar indicate directed interactions (irrespective of source), which are the ones shown on top of the chart.

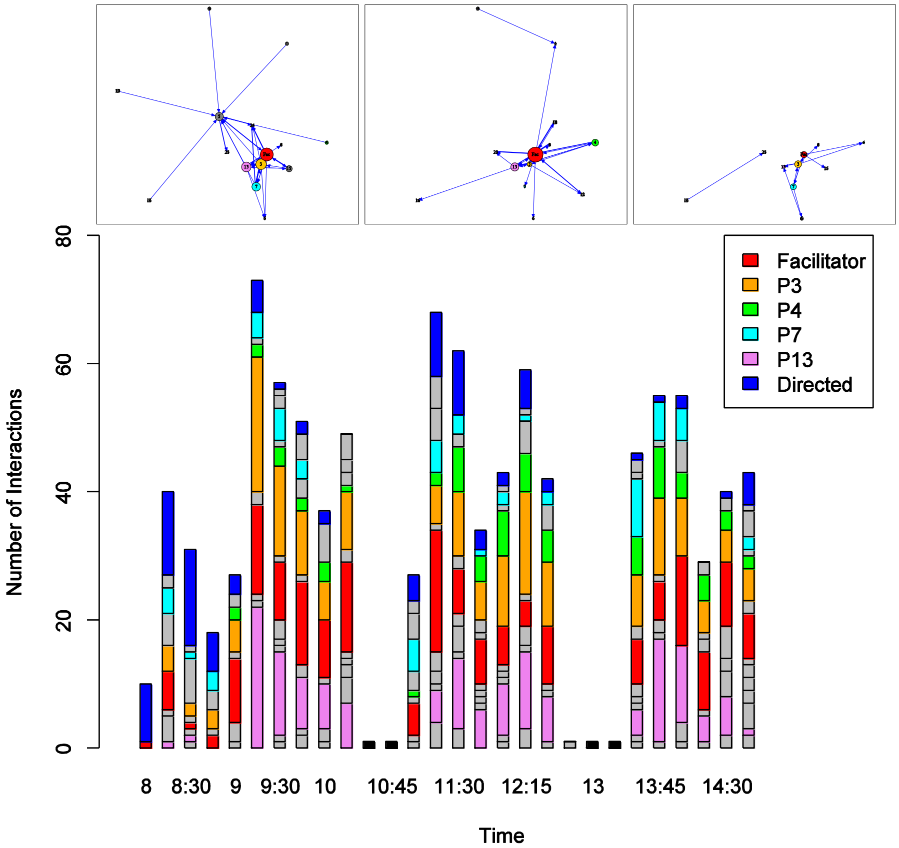

Figure 8. Bar chart of interactions (broadcasts colour-coded as in Figure 7, directed interactions dark blue). Bars show number of interactions for 15-minute periods. Breaks divide the workshop into 3 phases, corresponding to the network time slices at top.

The workshop was split by break periods into three distinct phases (8:00 to 10:30, 11:00 to 13:00, and 13:30 to 15:00). The three small network diagrams at the top of Figure 8 show corresponding time slices of the network of directed interactions. Directed interactions occurred mostly in the first two phases, and the pattern of interaction involving the senior stakeholder P5 occurred only in Phase 1. Each bar in Figure 8 represents a 15-minute interval. The taller bars therefore represent more dynamic periods of interaction, with more, but shorter, interactions. The periods from 9:15 to 9:30 and 11:15 to 11:45 are particularly notable for including many interactions.

As in our illustrative case study, a sliding-window plot of centrality offered considerable insight into the dynamics of the workshop. For the purpose of centrality calculation, broadcast communications to the whole group were treated as simultaneous directed interactions to all other participants. A 60-minute sliding time window was used for the centrality calculation. To avoid the problems of sliding windows overlapping break periods, we scaled centralities by dividing each score by the average centrality for the relevant time window. Figure 1 shows the result. Compared with the bar chart in Figure 8, Figure 1 gives a deeper insight into the interaction patterns during the course of the workshop, highlighting changes in "dominance" over time. Nine sub-phases are clearly visible:

Phase 1A: For a few minutes after 8:00 there is an initial dominance by the facilitator, setting the scene.

Phase 1B: Following this, and continuing to about 8:45, interaction was dominated by P5, the senior stakeholder. This portion of the workshop included many directed interactions.

Phase 1C: From about 8:45 to 9:00, the facilitator re-asserted control of the workshop.
Phase 1D: Following this, to about 10:00, there was a dynamic discussion dominated by P3 and P13.
Phase 1E: From about 10:00 to the break at 10:30, the facilitator again dominated interaction.
Phase 2A: After the break, from about 11:00 to 11:45, there was a continuation of Phase 1E, with the facilitator using directed interactions to elicit input from participants.
Phase 2B: Then, to the break at 13:00, P3 and P13 again dominated discussion, with input from P4.
Phase 3A: Phase 2B continued from about 13:30 to 14:00, with P7 also participating strongly.
Phase 3B: Then, to 15:00, the facilitator again re-asserted control, as discussion began to wind up.

5. CONCLUSION

In this paper, we have presented a number of techniques for temporal social network visualisation, and we have illustrated these techniques using an Internet discussion group and a government planning workshop. Our focus has been on two-dimensional visualisations, both with and without an explicit time axis. The former have included colour-coded bar charts showing the variation in interaction over time (Figure 3, Figure 8), as well as sequence diagrams. The most useful visualisations of this type have been centrality plots (Figure 4), particularly when scaled with respect to the average centrality (Figure 1). Such scaled centrality plots very clearly highlight changes in workshop dynamics over time, and therefore satisfy our goal of providing clear feedback to workshop facilitators on the overall “flow” of their workshop over time.

We have also considered temporal visualisations not involving an explicit time axis, specifically colour-coding of arrows (Figure 5) and multi-dimensional scaling of temporal activity. While interesting, these visualisations are more difficult to interpret, and hence less useful in providing feedback to workshop facilitators. However, we intend to explore their suitability for other tasks in future work.

The use of scaled centrality plots provides a useful addition to our toolkit for analysing workshop dynamics. Because the computation of centrality is relatively straightforward, the possibility exists of incorporating such visualisations into tools for recording workshop interactions, and therefore of providing real-time feedback on workshop dynamics.

REFERENCES

Bender-deMoll, S. and McFarland, D.A. (2006), “The Art and Science of Dynamic Network Visualization,” *Journal of Social Structure* **7** (2): www.cmu.edu/joss/content/articles/volume7/deMollMcFarland

Brandes, U. (2001), “Drawing on Physical Analogies,” in *Drawing Graphs: Methods and Models*, Kaufmann, M. and Wagner, D., eds., 71–86, Springer Verlag (Lecture Notes in Computer Science 2025).

Dekker, A.H. (2001), “Visualisation of Social Networks using CAVALIER,” Proceedings of the Australian Symposium on Information Visualisation, Sydney, Eades, P. and Pattison, T., eds. *Conferences in Research and Practice in Information Technology* **9**: crpit.com/confpapers/CRPITV9Dekker.pdf

Dekker, A.H. (2005), “Conceptual Distance in Social Network Analysis,” *Journal of Social Structure* **6** (3): www.cmu.edu/joss/content/articles/volume6/dekker

Dekker, A.H. (2008), “Centrality in Social Networks: Theoretical and Simulation Approaches,” *Proceedings of SimTecT 2008*, Melbourne, Australia, 33–38, Simulation Industry Association of Australia (ISBN: 0-9775257-4-0).

Fleming, C. (2007), “Exploring Social Complexity and Concept Evolution in War Games using Modified Dialogue Mapping,” Presentation to 8th Asia-Pacific Complex Systems Conference: www.complex07.org/online-edition/files/357.pdf

Fleming, C. (2008), personal communication.

Fowler, M. and Scott, K. (1997), *UML Distilled: Applying the Standard Object Modelling Language*, Addison-Wesley.

Freeman, L.C. (2000), “Visualizing Social Networks,” *Journal of Social Structure* **1** (1): www.cmu.edu/joss/content/articles/volume1/Freeman.html

Gloor, P. (2005), “Capturing Team Dynamics Through Temporal Social Surfaces,” Proceedings of 9th IEEE International Conference on Information Visualisation (IV’05).

Howison, J., Inoue, K., and Crowston, K. (2006), “Social Dynamics of Free and Open Source Team Communications,” *Proceedings of the IFIP 2nd International Conference on Open Source Software, Italy*.

Snijders, T.A.B. (2005), “Models for Longitudinal Network Data,” in *Models and Methods in Social Network Analysis*, Carrington, P.J., Scott, J., and Wasserman, S., eds., 215–247, Cambridge University Press.

Wasserman, S. and Faust, K. (1994), *Social Network Analysis: Methods and Applications*, Cambridge University Press.

Watts, D.J. (2003), *Six Degrees: The Science of a Connected Age*, William Heinemann.