Visualising Collaboration: Qualitative Analysis of an Email Visualisation Case Study

Onn Azraai Puade, Theodor G Wyeld
ITEE, The University of Queensland, Australia
Media, H&SS, The University of Adelaide, Australia
{puade@itee.uq.edu.au, theodor.wyeld@adelaide.edu.au}

Abstract

This paper reports on continuing work on visualising email collaboration [11]. It reports on the re-interviewing of participants of a previous email collaboration visualisation study regarding the identification of key players. Participants were asked to comment on the finding of key player impact on the collaboration as determined by our analysis and methods. We found, while they mostly agreed with our analysis they expressed reservations regarding the methods used. This forms the ground work for yet further work in developing a real-time visualisation tool for email-mediated collaboration. The qualitative analysis case-study method used in this study helped gain a deeper understanding of the nature and characteristic of the collaboration that would otherwise have been hidden in a quantitative analysis.

Keywords: Collaboration Visualisation, Email Visualisation, Case Study, Social Networks.

1. Introduction

This paper builds on work reported in an earlier paper in IV2006 on visualising email collaboration [see 11]. Email is a key collaboration medium for virtual teams. Email allows for asynchronous communication for exchange of information. The earlier paper was focused on identifying the key players in collaboration via email visualisation. Key players was identified as those with the greatest impact on the group as a whole. The second case study, reported in this paper, analysed the same collection of 176 emails over a period of 6 months of a collaborative activity. A one week, peak period, was isolated to a subset of 24 emails involving 10 participants. It is in these emails that we found the key players, their impact on the group, and how others felt about this after the event.

In the second study, analysis of the data revealed interaction patterns that occurred during the collaboration. While applying social network analysis techniques, such as who-talks-to-whom network graphs, we also investigated the content of each email and asked each participant to rate them according to importance. This was followed by a qualitative survey to clarify our interpretations of the reasoning behind their ratings. This case-study qualitative analysis helped us gain a deeper understanding of the nature and characteristic of the collaboration that would otherwise be hidden in a quantitative analysis. We found, while they mostly agreed with our analysis they expressed reservations regarding the methods used.

2. Collaboration Visualisation

Computer supported visual representation allows for different understandings of data [4, 13]. Collaboration in a virtual team can be described as the act of participants working together on a common task or process [3]. Visualisation of communication between team members allows for structural modelling and graphical representation of the interrelated elements.

Among the various computer-mediated communication technologies available to support collaboration, email communication is perhaps the most common [6, 15]. Divitini and Farshchian [5] describe email as the key collaboration medium, categorizing the many roles of email in collaboration. They claim it can be used to access experts, resolve issues and decisions, provide awareness to collaboration-related issues, and support irregular synchronous collaboration.

Several advantages of studying email as a measure of collaboration include its low cost and high volume nature, the social networks it forms, and the structural and temporal elements it automatically records [1]. These characteristics and attributes can be easily manipulated to construct different kinds of graphical representations of the communication data.

Visualisation of email communication patterns is not new. For example, Gloor et al's [8] collaborative innovation networks, Perer et al [10] and Viegas et al's [14] analyses of temporal rhythms of relationships and Samiei et al [12] and Kerr and Wilcox's [9] personal email management tool. Visualising email data assists information retrieval process and analyses of the trends embedded [7].

Perer et al [10] summarize the different types of interaction in email collections according to two

11th International Conference Information Visualization (IV'07)
0-7695-2900-3/07 $20.00 © 2007 IEEE

dimensions (see Table 1). Our study addresses the category in cell E. Adopting the socio-centric perspective, we looked at our collection of email messages as the raw data to help identify the many facets of the collaboration and help extract patterns of its structure.

Table 1 Types of interaction with email collections

	Individual	Organizational	Social
Current	Managing an individual user's current inbox (A)	Managing current email within an organization (B)	Managing current conversations within a public space (C)
Archived	Exploring an archive of an individual's message (D)	Exploring an archive of an organization's messages (E)	Exploring an archive of a public space (F)

3. Case Study 01 in brief

Case study 01 focused on email visualisation of a specific, time-constrained, event-driven collaboration. The collaboration in question was the organization of a workshop for a massive multi-user game resource development. The workshop ran for three days. 20 participants from 6 organizations were involved in the overall activity. The analysis period taken for the case study was a peak period just before, during, and after the workshop was run. 24 emails were identified, sent by 10 participants over this period, involving 11 participants from 3 different organizations.

3.1. Stakeholders

The stakeholders in that case study consisted of participants from diverse backgrounds. A group of 8 males and 3 females, they represented different functional roles in the collaboration: facilities manager, project coordinator, technical assistant, chief executive officer, secretary, research assistant, external consultant, project leader, artist, programmer and project manager. They represent a wide range of ages 21-51. Their acculturation to email as a communication tool was assumed.

3.2. Process

The contents of all 24 emails were analysed. The connection between all the participants was plotted to identify any collaboration structural features. A survey was then conducted with the same participants. Each of the 10 participants was given a printed copy of all 24 emails arranged in chronological order. Participants were asked to rate each of the emails in terms of importance on a scale of 0 (Not rated), 1 (Not important), 2 (Important), and 3 (Very important). A comments area was included for them to comment and provide reasonings for each rated email.

Investigating the reasonings given by the participants in the comments fields, we identified the various approaches taken by the participants when rating the emails and common trends across all email ratings rationales. A follow-up survey was conducted to solicit their opinions on our findings.

3.3. Analyses

Network graphs were constructed using Pajek [2], a social network analysis visualisation tool. Undirected graphs, with vertices representing emails and nodes representing participants were used. This structural representation helped identify how all 11 participants were connected, and how it might resemble the actual organizational structure of the collaboration [see 11].

Data from the initial survey was analysed and ratings for each of the emails were tabulated. This information was organized according to the number of emails per participants, ratings and average ratings. From this organization strategy, the average ratings were used to represent the 'loudness' (L) of a participant's message in the collaboration. This was then multiplied by the number of emails (N) they sent to get the measure of their overall 'impact' (I) on the collaboration ($L \times N = I$). The result identified participant 'CG' as having the highest impact factor in the collaboration.

This information was then used to create a visualisation schema. We created a proof-of-concept visualisation to allow us to gain an understanding of the interrelationships of the various data analysed.

4. Case Study 02

Some months after the first case study results were made known to the participants, we re-interviewed the same participants to solicit their responses to these results. We found most agreed with our findings but for many different reasons. This level of information feedback is important in evaluating the potential of this system as a real-time application for further research. Of the 11 original participants, only 6 were available for the follow-up survey.

4.1. Responses to Case Study 01 Findings

A survey was conducted with the same participants from the first case study. In it, we asked them to comment on the finding that participant 'CG' emerged as the key player. We then analysed, in detail, their responses to this question. We compared this with their reasoning for rating the original 24 emails in the first case study.

We found there were three main types of responses identifiable by the different approaches each participant adopted when they rated each email and provided their reasonings. These included: Authoritative (A), Action-o-oriented (AO) and Non-committal (NC) (see Table 2).

In the follow up survey, participants were asked how they 'felt' about 'CG' being identified as the key player.

None expressed any surprise at this result, although some felt it was only due to others not being available at the time to fill this role (see Table 3).

Table 2 Different types of response made by the participants

Response	*Description*
Authoritative (A)	Rated the email according to how they saw their position, rank, official function or purpose in the collaboration
Action-oriented (AO)	Positively described what was referred to in the email by identifying it as an action item
Non-committal (NC)	Rated and responded with non-commitment, often negatively

Table 4 The actual different responses on 'CG' being the key player

Ind.	*Feedback*
JH	CG was very engaged and interested and took a leadership position.
BL	I think this is accurate and true reflection.
AK	I think that as the "action" person of the group that would be expected.
CG	I was in a position to work between the pure technical development and the artistic development. I tried to encourage artistic development and deliver the appropriate materials to the technical people. My emails were directed in this effect, as well as trying to keep a positive energy/momentum going.
TW	I am not surprised that CG emerged as the key player because he was very active in the group and jumped in and got things organized when others didn't or wouldn't.
JC	Due to other production deadlines at the time, CG was the person with the most time to converse with the other people. CG was also the perfect person for the job.

Table 4 The actual different responses on the method used

Ind.	*Feedback*
JH	May have problem if a person is a relay role, otherwise ok.
BL	I think this is simplistic and there are other factors, time, personality, situation, context, attitude, etc.
AK	I think this is a reasonable and logical method of measurement.
CG	I'd be interested also on what average rating divided by number of emails. Then compare this to weighing mentioned above. Also importance by timeline? Is there a pattern when compared to other people through project timeline?
TW	I think this seems ok but I can't be sure if it is really accurate.
JC	Using emails as an indicator can work for the beginning stages of the collaboration but doesn't cater for the intense time when we came together

The same participants were also asked to comment on how they felt about our 'methods' for arriving at the conclusion that 'CG' was identified as the key player. Their responses to this question were more varied than those referring to the results *per se* (Table 4).

5. Discussion

In case study 01, we were able to identify the connectivity between participants in the collaboration by using network graphs to represent communication patterns. This allowed us to identify highly connected clusters of participants, which we could isolated as key participants. The email rating on the other hand, allows for weighting of participants, better illustrating the importance of these key participants. In this way the connectivity visualisation was extended to include how important emails and participants were perceived by others. Ratings could also be aggregated and averaged for comparison purposes.

As the first case study survey required participants to rate emails in terms of importance, we could find how important a participant's email was rather than focusing merely on the number of emails sent. We considered the average rating applied similarly to all participants as a measure of loudness; how loud his or her messages were received by others. From this we found how important a participant's contribution was as a factor of his/her loudness and how many emails he/she sent as a measure of impact on the overall collaboration. We applied these calculations and created a visualisation schema to represent the data.

In the second case study, looking at the ratings and reasoning given by the same participants, we identified three different types of response. Participants by response can be graphically represented as falling into or across those three types (see Figure 1).

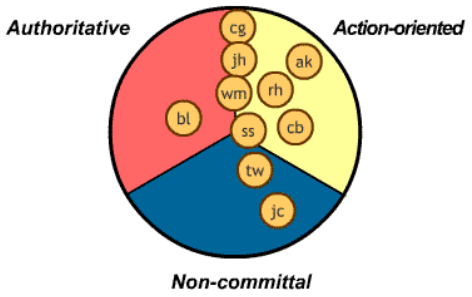

Figure 1: Participants by type of ratings given

A collaboration team member who rated emails based on an authoritative approach tended to rate emails according to their perceived position, rank, official function or purpose in the collaboration. For example, the project manager (JH) rated half of the emails in an authoritative manner. Characteristically, their comments were described in terms of their function in the collaboration group. Such as, this email refers to "passing the progress report to the key team members",

or that email is of the "type communication/clarification of approval". Those with an authoritative response tended to rate emails directly related to the management of the collaboration as important or neutral.

A collaboration team member who rated emails based on an action-oriented approach tended to positively described what was referred to in the email. They rated the emails by identifying occurrences in the email that were important, or not important, to the collaboration. Whether they actually had received the email or not, or whether they were directly involved in the communication or not, was not important. They tended to infer meaning on the email as part of an ongoing reflective narrative. For example, the research assistant (CG) included comments such as, this email refers to "confirmation of meeting time and place", or "describe[ing] movement of project leader". Action-oriented raters tended to rate most emails highly.

A collaboration team member who rated emails based on a non-committal approach tended to be more negative about emails. For example, either emails were "not interesting" or they were "not my problem". Non-committal raters tend to rate emails lowly. However, when an email directly involved them, they rated it more highly.

In the second part of the second case study survey, we asked the same participants how they felt about our findings that 'CG' emerged as the key player. All agreed with our findings. This is despite the option to argue against this finding and nominating someone else as the key player. From their comments, they described action and active leadership as the key indicator of key players in the collaboration. CG was reported as displaying both these attributes in the collaboration, further underscoring the numerical results.

The attribute values from the second study are sorted by participant in Table 6. The first column lists participants and their role. The second column contains descriptions of how each participant rated emails and how this relates to their role. The third column shows number of emails that they sent during the peak period being studied. The fourth column shows the average rating they received from others for the emails they sent – not including their own rating. The fifth column shows their overall impact factor when all participants and all email ratings are taken into account. The sixth column shows the overall average rating that each participant gave to all emails.

Table 6 Table showing participants, response type in email rating, average rating, impact factor, and average rating given in the case study

Participant	*Response type in rating email*	*Number of emails sent*	*Average rating received by others*	*Overall impact factor*	*Average rating for all emails by all participants*
CG – research assistant	CG email ratings were balanced between action-oriented (AO) and authoritative (A). He positively described the occurrences throughout the email collection, and rated action-oriented emails with high ratings.	5	1.9	9.5	2
TW – coordinator	TW email ratings were a mixture of all types of response: average authoritative (A) and action-oriented (AO) but high non-committal (NC) responses. He rated emails related to him highly, and others as "not interesting". This is the typical role for a coordinator: focusing only on matters that he needed to look at.	5	1.84	9.2	1.71
JH – project manager	JH ratings were Authoritative (A) and action-oriented (AO). This is typical of a management role. Emails with process and progress information were rated equally important.	1	1.9	1.9	1.71
AK – secretary	AK ratings were action-oriented (AO). In a secretarial role, she rated communication emails as important and others' personal emails as not important.	1	2	2	1.79
BL – project manager	BL ratings were non-committal (NC) and authoritative (A). Emails related to him were rated highly, while other were rated more lowly or not rated at all.	2	1.65	3.3	1.63

From Table 6, we can say that the fact that CG rated all emails highly and TW rated his own highly tended boost their overall impact. This is due also to the sheer number of emails they sent. This is reflected also in the ratings that BL gave for emails related to himself. AK and JH, on the other hand, rated all emails including those related to them in an average manner, thus also achieving an average rating for themselves.

Although it is difficult to make any definitive direct numerical correlation, due to the small group involved in this second study, when we look at their rationales for why they rated emails the way they did, there seems to be some correlation between this and their overall impact. Whether this is due to the nature of their response type (see Figure 1), or the case study environment, is not clear. What is clear is that what they said and what others said about them seems to fit our numerical analysis. Therefore, further work on implementing a real-time application to track

collaboration in action, albeit a more fully-featured version, is worth pursuing.

6. Conclusion

Case study 01 explored visualising collaboration. Email was identified as a key collaboration medium. A subset of emails collected from a collaboration activity was used to help understand the various aspects of collaboration by way of several different visualisation analyses. Rather than simply analyse the number of emails sent by who to whom, we analysed the detailed contents of these emails. Using the email ratings as a basis, we identified the key player in the collaboration by calculating and comparing the impact factor for all participants.

A survey was used to solicit responses from the same participants to the first study results. All the participants involved in the second study agreed with our numerical findings. They further described 'action' and 'active leadership' as indicators for key players in collaboration. From the qualitative analysis, we also identified the different approaches taken by participants when rating emails. Each participant's approach was correlated with his or her position or role within the collaboration.

Despite some shortcomings (not all participants participated in the second case study) the insights gained through the qualitative approach used in these studies allowed for a more comprehensive understanding of the collaboration and the participants involved. This detailed analysis of rationales is often ignored, discarded, or missing from the more typical quantitative analysis. Our findings and the participants' suggestions will be taken into account in the next phase of this project – the development of a visualisation tool to track collaboration in real-time. It is envisaged that future implementation of such a tool could be used to foster awareness of participant roles within dynamic collaborations.

7. Future directions

In both case studies, we looked at how participants rated emails and compared their ratings with the different reasons given. This helps us understand how each participant perceived others' importance to the collaboration. While they tended to agree with our identification of 'CG' as the key player, they suggested several interrelated factors that could have affected the perceived importance of a single email or participant. They would like to see this reflected in any future implementation of a real-time email collaboration visualisation application. Other features discussed include: timeframe, response type, context, and attitude.

References

[1] M. V. Alstyne and J. Zhang, "EmailNet - A System for Mining Social Influence and Network Topology in Communication," 2003.

[2] V. Batagelj and A. Mrvar, "Pajek - Program for Large Network Analysis," Connections, vol. 21, pp. 47-57, 1998.

[3] R. P. Biuk-Aghai, "Visualization of Web-Based Workspace Structures " in Proceedings of the First International Conference on Web Information Systems Engineering (WISE'00)-Volume 1 - Volume 1 IEEE Computer Society, 2000 pp. 302

[4] S. K. Card, J. D. Mackinlay, and B. Shneiderman, Readings in information visualization: using vision to think. San Francisco: Morgan Kaufmann Publishers, 1999.

[5] M. Divitini and B. A. Farshchian, "Collaboration and Coordination through Basic Internet Tools: A Case Study," presented at World MultiConference SCI/ISA, Orlando, Florida, 1999.

[6] N. Ducheneaut and V. Belotti, "Email as a habitat: An exploration of embedded personal information management," ACM Interactions, vol. 8, pp. 30-38, 2001.

[7] S. Frau, J. C. Roberts, and N. Boukhelifa, "Dynamic Coordinated Email Visualization," presented at WSCG 2005, Plzen, Czech Republic, 2005.

[8] P. A. Gloor, R. Laubacher, S. B. C. Dynes, and Y. Zhao, "Visualization of Communication Patterns in Collaborative Innovation Networks - Analysis of Some W3C Working Groups " in Proceedings of the twelfth international conference on Information and knowledge management New Orleans, LA, USA ACM Press, 2003 pp. 56-60

[9] B, Kerr, and E. Wilcox, "Designing reMail: reinventing the email client through innovation and integration," paper presented at the Conference on Human Factors in Computing Systems, Vienna, Austria, 2004.

[10] A. Perer, B. Shneiderman, and D. W. Oard, "Using Rhythms of Relationships to Understand Email Archives," Journal of the American Society for Information Science and Technology, 2005.

[11] O. A. Puade, and T. G. Wyeld, "Visualising Collaboration via Email: Finding the Key Players," in Information Visualization 2006, pp. 124-129, 2006.

[12] M. Samie, J. Dill, and A. Kirkpatrick, "EzMail: Using Information Visualization Techniques to Help Manage Email," paper presented at the Eighth International Conference on Information Visualization, 2004.

[13] B. Shneiderman, "The eyes have it: a task by data type taxonomy for information visualizations," presented at Visual Languages, 1996. Proceedings., IEEE Symposium on, 1996.

[14] F. B. Viégas, D. Boyd, D. H. Ngunyen, J. Potter, and J. Donath, "Digital Artifacts for Remembering and Storytelling: PostHistory and Social Network Fragments," presented at Proceedings of the 37th Annual Hawaii International Conference on System Sciences, 2004.

[15] S. Whittaker and C. Sidner, "Email overload: exploring personal information management of email," presented at Proceedings of the SIGCHI conference on Human factors in computing system, Vancouver, British Columbia, Canada, 1996.