

Collaborative Visualization for Supporting the Analysis of Mobile Device Data

Thomas Ludwig, Tino Hilbert and Volkmar Pipek

Abstract Visualizations are mainly used for providing easy access to complex information and data. Within this paper we focus on how visualization itself can serve as a collaborative aspect within distributed and asynchronous team work. In doing so, we try to uncover challenges to support a team of researchers in understanding and analyzing mobile data by collaborative visualization. Based on a review of recent literature, two workshops with participants from the academic field were conducted, which revealed use cases and major design challenges for a collaborative visualization approach. With our user-centered study, we introduce design implications for collaborative visualizations that focus on research questions instead on single visualizations, embed multiple visualizations into a discussion thread, highlight relations between research artefacts as well as include external parties in collaborative visualizations.

Introduction

In recent years, smart mobile devices have become continuously more ubiquitous in daily life. Equipped with low cost sensors such as accelerometers and gyroscopes, they allow data to be collected in situ, remotely and parallel from multiple devices (Hagen et al. 2007). Based on the increasing usage of one device for both private and professional work purposes, smart mobile devices therefore offer

T. Ludwig (✉) · T. Hilbert · V. Pipek
Institute for Information Systems, University of Siegen, Siegen, Germany
e-mail: thomas.ludwig@uni-siegen.de

T. Hilbert
e-mail: tino.hilbert@uni-siegen.de

V. Pipek
e-mail: volkmar.pipek@uni-siegen.de

a great opportunity to create a holistic view of the users' appropriation of such devices. Approaches already exist, which try to cover various aspects related to the mobile sensing of gathering data or mobile data mining to discover hidden usage patterns. Beyond data gathering, there are a variety of methods for analyzing mobile-gathered data. Those methods encompass mainly highly collaborative tasks including teams and multiple members with different backgrounds, located at different places and working at different times. CSCW as an interdisciplinary field influenced by various communities from computer science, social science, and psychology is conceived as 'an endeavor to understand the nature and characteristics of cooperative work with the objective of designing adequate computer-based technologies' (Bannon and Schmidt 1991). By creating an understanding of how multiple actors work in different collaborative settings, CSCW explores how groups in an organization can be supported by tools and how these tools might change the organization itself and vice versa. An elementary aspect is awareness that builds the base for exploration, evaluation and design of supportive methods/tools within work context (Bannon and Schmidt 1991).

One important field of mobile-gathered data analysis is 'information visualization', whereby visualization of data and information is not limited to explanation purposes. Instead, the visualization itself can be an independent approach to explore and analyze data based on cognitive and perceptual principles. Within this paper we try to combine the research area of CSCW with concepts of visualization and examine when and how 'collaborative visualization' can be used to support a team of different researchers in understanding and analyzing mobile-gathered data. In doing so, we review recent literature regarding collaborative visualization and its origins 'visualization' and 'CSCW'. Based on the identified design challenges, we conducted two workshops with actors from the academic field where possible use cases for collaborative visualizations and design challenges for analyzing mobile-gathered data are identified. The basis of our approach is a previously developed mobile application which utilizes the concept of participatory sensing to gather mobile data for research projects (Ludwig and Scholl 2014).

Related Work: Collaborative Visualization

Visualizations are not a modern day invention. Cartography and astronomy have been using visual representations since 200 B.C. and the 10th-century respectively. Visualizations and graphics are used for a wide range of fields even beyond research contexts: for instance in journalism, to provide a broad audience with easy access to complex information. Computer science provides various approaches and tools for gathering, processing and analyzing huge amounts of data. Chen et al. (2009) describe different processes of how interactive visualizations are created and how they can be supported by existing information. Data visualization is 'the use of computer-supported, interactive, visual representations of data to amplify cognition' (Card et al. 1999) and can be subdivided into

information visualization and scientific visualization, whereby scientific visualization focuses on physically-based, scientific data and information visualization on abstract, non-physically-based data (Card et al. 1999). Collaborative visualization can be understood as “the shared use of computer supported, (interactive), visual representations of data by more than one person with the common goal of contribution to joint information processing activities” (Isenberg et al. 2011). The idea behind collaborative visualization resulted from the need to overcome the traditional design of single-user visualization systems and to allow the collective exploration and analysis of large data sets through visualization.

Current Approaches of Collaborative Visualization

First approaches of collaborative visualizations were redefined extensions of existing modular systems for collaborative use cases. These were mainly achieved by duplicating views or sharing some selected parts—or a mixture of both approaches (Wood et al. 1997). Former research has shown that visualizations are significant for collaborative work, e.g. the benefits of using visualizations compared to not using them (Bresciani and Eppler 2009); or that groups obtain better results with visualization systems in comparison to individuals (Mark et al. 2002). Isenberg et al. (2011) differentiate collaborative visualization systems into the two categories of distributed and co-located approaches. In both cases, various approaches exist (Isenberg and Carpendale 2007).

Hugin (Kim et al. 2010) is a mixed-presence tool which supports co-located as well as distributed collaboration in a synchronous working context and examines coordination mechanisms of awareness, territories and access control. For highly distributed and asynchronous settings, such as those of researcher teams, we focus on asynchronous and distributed approaches. As Willett et al. (2011) mentioned asynchronous collaboration is often based on the decomposition of work into smaller tasks which can be performed in parallel. Thus, central mechanisms of such tools are based on creating awareness and aggregating individual results. Nevertheless, for a better understanding of visualizations and their applications, knowledge about cognitive and perceptual principles is necessary.

ManyEyes (Viegas et al. 2007) is a tool which supports the asynchronous collaborative analysis of social data. ManyEyes is public and can be accessed by anyone who is interested. It maps the visualization process, beginning with the upload of data up to the discussions about the visualizations. Users can upload data sets and are able to create visualizations. ManyEyes provides various kinds of visualizations, like bubble charts or network diagrams, and supports the visualizations with collaborative features such as annotations or feedback and discussions prompted by comments. To obtain an in-depth understanding of how ManyEyes is used as a community, interviews with users and existing logs were analyzed after the public launch. It was shown that ManyEyes is not used as a dedicated community, but as a platform for other, external communities where large proportion

of the communication takes place off-site (Danis et al. 2008). Heer et al. (2009) present another approach for the collaborative visual analysis of data with Sense.us. In contrast to ManyEyes users do not have the possibility to upload their own data sets. Sense.us is an analytics tool that supports view sharing, doubly linked discussions and social navigation mechanisms, e.g. listings of comments and recent activities. Heer et al. (2009) describe how central mechanisms, such as providing collaborating participants with access to the same visual environment; using graphical annotations to refer to direct conversations; or separating annotations and comments visually from the visualization, were used by participants and determined that it was especially the combination of mechanisms which allowed a deep exploration of the data. CommentSpace (Willett et al. 2011), which is partially based on the experiences previously gathered by Sense.us and ManyEyes, focuses especially on the discussions around visualizations. Willett et al. (2011) criticized that the previous approaches did not support their users in more complex analytical tasks, e.g. the gathering of evidence. CommentSpace therefore focuses on adding a small and fixed vocabulary for tagging and linking to comments. Such mechanism is especially designed to support the generation of hypothesis and the gathering of evidence. This is done by using tags like 'question', 'hypothesis' or 'to-do' and links such as 'evidence-for' or 'evidence-against'.

The presented approaches have shown that tagging and linking, if used, can have a positive impact on the collaborative analysis tasks. During deployments it was noted that tags and links were not used as often as in laboratory evaluations. Willett et al. (2011) assume motivational factors and emphasize the need for guidance and incentives to facilitate the benefits of the provided mechanisms. Based on the understanding of both fields of visualization and CSCW, as well as their implications for collaborative visualizations, critical issues for collaborative visualization designs were identified.

Design Issues

Heer and Agrawala (2008) identified important design considerations for collaborative visual analytics, which were based on prior experience of Sense.us (Heer et al. 2009) and a survey of relevant research areas including CSCW and visual analytics. In the course of their work, they identified seven areas: (1) Division and allocation of work; (2) Common ground and awareness; (3) Reference and deixis; (4) Incentives and engagement; (5) Identity, trust and reputation; (6) Group dynamics, and (7) Consensus and decision making. For these areas they went on to present 24 design implications for asynchronous collaborative visualization systems, e.g. 'Artefact histories' as part of the area 'Common ground and awareness'; or 'Personal relevance' for the area 'Incentives and engagement'.

Isenberg et al. (2008) analyzed how individuals and teams work and interact during visual information analysis tasks. Based on observations of individuals and teams, who had to solve tasks in a co-located and synchronous setting,

Isenberg et al. (2008) identified eight different processes: During the *Browse process* the team scans through the available data and artefacts to form their first impressions of the available data. The *Parse process* involves the (re-)reading of the task to create a common understanding of the problem and how to solve it. During the *Discuss Collaboration Style process* the team discusses the overall task division strategy. In the *Establish Task Strategy process*, the team figures out the best way to perform the tasks with the available data and tools. The *Clarify process* involves activities which help to understand information artefacts. The *Select process* is about finding and selecting relevant information artefacts for a particular task. The *Operate process* includes higher-level cognitive work on a specific view of the data to extract information for fulfilling a task. During the *Validate process* the team tries to confirm the solution of a task. It also involves activities which ensure that the process of the team itself is correct. Isenberg et al. (2008) mentioned that the temporal order of the process differed from team to team and one typical temporal order does not exist. For a collaborative visualization, three general implications must be addressed: It must be flexible regarding the temporal sequence of work processes: it must support changing work strategies, and it must support workspaces.

Research Approach

Currently, existing asynchronous distributed approaches for collaborative visualization like ManyEyes, Sense.us or CommentSpace are all based on design hypotheses which were previously formulated. As is usual within the discourses of Information Systems, design assumptions were derived from literature. From a CSCW and more practice-oriented perspective, we argue that there are several shortcomings stemming from the existing approaches:

1. It is questionable how the design decisions can meet the user requirements without involving the users themselves. Feedback from users which is only considered after the deployment means that it is often influenced by the capabilities and limitations of the tool.
2. None of the tools presented were designed to be used in the context of a group sharing the same goal, as it was the case regarding researcher teams. In contrast, ManyEyes in particular was designed for an unspecified and broad audience. This causes significant differences, for e.g. privacy issues or data structure.
3. None of them mapped the visualization process starting with the gathering of data and its exploration, and ending with the analysis of data. Especially within long-term studies, data explorations often start before the collection is finished. ManyEyes provides uploading final data but does not involve aspects about its source. Sense.us does not even provide the possibility to upload one's own data.

Within this paper we therefore focus on a more user-centered approach including ideas from potential users for applying collaborative visualization for analyzing mobile-gathered data. The underlying purpose is to create a comprehensive picture of relevant context of use. The resulting concept should support researchers during exploration and analysis of data gathered by mobile devices. We focus on long-term studies, where it is possible to explore data even before gathering is finished. This ‘open heart empiricism’ provides the option to gain first insights into the data while it is being collected, and this can then be used to adapt the ongoing research strategy.

Workshop

To gain a deeper understanding of how collaborative analysis tasks could be supported by visualizations, we conducted two workshops. The first workshop lasted three hours and consisted of six participants; one female and five male. Their age ranged from 25 to 29 years. Five of them were research associates and one was a research student. All had several years’ experience in an academic context. Their context of research differs and covers research fields ranging from mobility to crisis management. None of the participants had any previous experience with collaborative visualization. The workshop was split into two parts: The first consisting of an introduction, brainstorming and discussion of previous experiences. This served as an introduction to the topic and assisted the participants to familiarize with the context, especially the use and analysis of mobile device data within research projects. Brainstorming was divided into two phases, whereby each was led by its own central question: ‘What kind of data can you gather using personal mobile devices?’ and ‘What kind of research questions could you answer by using the data gathered by mobile devices?’ The intention was that the results should include different views from different kinds of project. The second part was intended to gather insights and design implications for possible IT support which aims at utilizing visualizations to support research teams in analyzing mobile data. It started with a presentation of ManyEyes and the participants had the opportunity to form their first impressions on ‘how collaborative visualization works’. After the presentation, a Brainwriting Pool activity was conducted with the central question of ‘which ideas and requirements do you have for approaches based on visualizations that support your team during a research project?’ (Figs. 1 and 2). The analysis was based on the main question: ‘Which design challenges need to be addressed for collaborative visualization?’ The analysis revealed various design issues, which were separated into the three areas *Visualization*, *Collaboration* and *Sharing* (Table 1). The second workshop was used to evaluate and redefine the raised issues by other participants from academia. Seven researchers participated in the second workshop. In the following, we present mainly the results of the first workshop, which were confirmed by the participants of the second workshop.

Table 1 Design issues

Design issue	
V.1	Select relevant data
V.2	Filter and aggregate data
V.3	Link data
C.1	Merge results
C.2	Provide structure
C.3	Encourage communication
C.4	Reveal relations
S.1	Build commitment
S.2	Allow participation

Fig. 1 Participant reviewing Brainwriting cards



Fig. 2 All Brainwriting cards



Visualization

Visualization includes design issues which are directly related to the visualization of the data mobile-gathered by researchers. Thus related questions are: ‘How can one visualize the data?’ or ‘Which capabilities and features regarding data and visualization need to be provided?’ In respect of this, three issues have been identified and outlined below.

The design issue ‘Select relevant data’ appears in various notes and comments. The question behind this issue is: ‘Which relevant data does the researcher need to visualize gaining the best possible insights into his/her research interests?’ The

initial note was *“I would like to visualize relevant locations”*. ‘Select relevant data’ was mainly discussed in context of the dimension ‘location’. That may have been caused by the participants’ research focus, but beyond this, the same questions can be transferred to most other categories of data. As implied by the participants, relevance depends on various aspects and differs due to individual factors. Supporting researchers in this issue may help during the selection of data for the visualization. However it should also provide an opportunity to focus on important elements of the resulting visualization.

The second design issue is ‘Filter and aggregate data’. The participants continually mentioned filtering and aggregation of data as being important capabilities which IT support should provide in context of research projects. Aggregation affords different levels of abstraction and enables researchers to explore a research problem from different angles. This was highlighted by one participant during the prioritization of the Brainwriting Pool: *“What are relevant places? What are relevant buildings? Or what may be relevant... well, that has got different levels, different levels of abstraction, which are based on the same data”*. Advantages of the data filtering are mentioned as they offer support in the context of exploring visualizations and their meaning by experimenting with different filters as well as using them as bookmarks and, therefore, providing options to share a current visualization.

The last design issue is ‘Link data’. The participants had several ideas regarding the linkage or enrichment of the mobile-gathered data. Notably, the opportunity to link visualizations with other data sources which provide additional information about the context of use was one reoccurring aspect for the participants: *“I would like to be able to extend data visualizations with qualitative data, e.g. questionnaires”* or *“How can I combine sensors with qualitative or other quantitative data? How can I get feedback from the users about their context?”* Remarkably, the participants focused mostly on capabilities related to the underlying data and possible operations based on them. Various aspects of the issues ‘Filter and aggregate data’ and ‘Link data’ were often mentioned and prioritized during the Brainwriting Pool. According to this, these two issues and their implications must be considered particularly when designing a possible IT support.

Collaboration

Collaboration includes issues related to features regarding team work. The central question is: ‘How can possible IT support be designed and structured to support the work of a research team beyond the process of visualization?’ Here our specific focus is on the actual research team; external stakeholders such as study participants who contribute their mobile data, or other involved stakeholders are covered in the section ‘Sharing’.

The first design issue is ‘Merge results’. It became clear that it is important to be able to combine or to merge visualizations and (interim) results with each other. It was stated that these combinations may be difficult to implement, but regarding

the merging of multiple data analysis, one of the participants said: *“Nevertheless, [merging multiple data analysis] does have great potential, as one is just able to reproduce several, more complex scenarios”*. As a solution, the participants suggested switchable layers.

The ‘Provide structure’ is another important design aspect: *“I need a good underlying workflow which structures the collaboration”*. Based on this comment, several suggestions were added, e.g.: *“One should be able to filter the workflow based on time or data”*. Additionally, related questions were raised: *“How can cooperation be structured? How can rights be managed?”* Furthermore, version control and change history were often mentioned as necessary functionalities. Keeping the history of visualization changes also provides other advantages: *“It helps team members to stay up-to-date, to prove validity and reliability of the former process and its results”* and it could also be used to support new team members to become acquainted with the project; or to learn by reviewing the previous work of other team members: *“... if you integrate new people at a later stage, the data changes are visible. Thus, s/he can image how the project evolved”*.

The design issue ‘Encourage communication’ includes the thoughts and ideas which refer to the capabilities related to the question: How can people be made aware of each other and their work; and how can communication be stimulated and supported by adequate features? The need to encourage communication was noted: *“The stimulation of discourses and group discussions (forums) for specific topics, their aims [and] the implementation of aims”*. Regarding this issue and especially the associated aspect of awareness, it was suggested that team members should be visible during shared editing, to provide user profiles and also to visualize the relationships involved in the cooperation. One of the participants suggested *“the creation of filters and data sets that can be shared”*. The motivation behind this idea was that if s/he is stuck in a vast amount of data, s/he could share the current filter and a team member could try to help. Additionally, this provides another opportunity for learning, as exemplified: *“This should allow me [to support other people] who work with similar data sets; telling them: ‘I performed an analysis, which might fit [your case]. Just have a look how I did it’”*.

The last design issue is ‘Reveal relations’. In the workshop, the participants several times highlighted relations between different entities of possible IT support: *“Filters can and should be added to certain research questions”* or *“Goals should be added to visualizations, if applicable, to make the collaboration more efficient”*. Additionally, they proposed linking related visualizations with similar goals as well as research questions respectively, to allow the addition of qualitative codes for visualizations. Regarding goals and visualizations, one participant stated: *“I think mostly [goals] are mostly defined by the project. But when I think of our group, goals change every six month. You keep getting new insights the whole time and you build [your future research] on them. You develop new goals and directions that must be addressed”*.

The issues presented by ‘Collaboration’ show that, in addition to capabilities for the visualization itself, IT support has to incorporate adequate capabilities for supporting collaborative work into its design. Both the issues ‘Provide structure’

and ‘Encourage communication’ have been considered, but in particular both revealed ideas and thoughts with a slightly different focus compared to the existing systems, e.g. Hugin or ManyEyes. The collaborative analysis and exploration of data is less focused on visualization and the discussion based on them. The considerations to highlight relations between projects, visualizations and goals provide an interesting new approach. The thoughts behind ‘Merge results’ point out that the participants do not consider that a single visualization provides the basis for visual analysis. In contrast, the visual analysis is understood to be a process that includes the merging and combination of multiple data analyses and results.

Sharing

The third area, ‘Sharing’, includes design issues which address the sharing of information and results with external stakeholders as well as their involvement in the collaborative analysis. External actors can be, for instance, participants, who contribute mobile data.

The first design issue, ‘Build commitment’, pertains to the considerations necessary to ensure the participants commitment to the study. One participant asked: *“How can I integrate people who supply the data so that sustained success is ensured?”* Because: *“If someone provides data, you have to give him something in return”*. Besides providing incentives like *“I would like to send visualized results to the participants”*, the researchers suggest allowing the participants to perform all data analyses themselves.

‘Allow participation’ is the second design issue. The participants suggested: *“The integration of feedback possibilities for participants”* and *“the creation of participation possibilities, so that participants are able to create visual analyses themselves”*. As shown, it is related to the previously mentioned design issue which utilizes participation as a possibility to promote commitment. The notes showed that feedback can also be used by the researchers to involve participants later on in the study or project phases. For instance, one participant explained: *“In context of research data sensing and analysis, one has to give the user something in return. So that they know progress has been made and, based on that, new goals can be developed collaboratively with the user”*. Notably, prior to the workshop, the importance of those issues were not expected to be as high. However, with regard to our approach, sharing and participation are not directly linked to visualization itself, but, in fact, crucial to a lasting success in the context of this work.

Conclusion

Due to the workshops we gained insights into the needs and thoughts of researchers regarding the design issues pertaining to the support of the analysis of mobile-gathered data through collaborative visualization. The identification of the three

areas outlined underlines the variety of issues which have to be considered when designing possible IT research infrastructures. An approach based on these insights might provide a valuable basis for computer-supported cooperative work in the context of visual analytics as further steps towards eScience. As expected, the participants provided new insights and ideas which had not previously been considered in literature. Related to the gathering of mobile data within long-term studies with participants, the issues ‘Build commitment’ and ‘Allow participation’ were highlighted as they not only state which requirements have to be met for lasting success; they also show which synergies can be created through the involvement and participation of data suppliers. Beyond that, further ideas were revealed. We found that, instead of focusing on groups that are formally defined as an organization, a tool of collaborative visualization within the academic context must focus on informally defined teams grouped by their interest in shared research questions and common goals. For example, some of the participants’ thoughts shifted the focus of collaborative analysis to underlying research questions and the relations between the different artefacts of such an approach. This has only been lightly addressed by existing approaches like CommentSpace (Willett et al. 2011), but not as deeply as the participants mentioned. This paper contributes first steps of using a CSCW lens towards existing approaches of collaborative visualization; and, based on the two workshops, we derived user-centered design issues for visualizations that encompass a collaborative analysis of mobile-gathered data and therefore contribute to eScience. As our next step we are currently implementing a web-based application building upon our uncovered issues for providing a tool to analyze mobile-gathered data aiming at fostering the collaboration between academic project teams and external stakeholders by simple visualizations.

References

- Bannon, L. J., & Schmidt, K. (1991). *Studies in computer supported cooperative work* (pp. 3–16). Amsterdam: North-Holland Publishing Co.
- Bresciani, S., & Eppler, M. (2009). The benefits of synchronous collaborative information visualization: Evidence from an experimental evaluation. *IEEE Transactions on Visualization and Computer Graphics*, 15(6), 1073–1080.
- Card, S. K., Mackinlay, J. D., & Shneiderman, B. (Eds.). (1999). *Readings in information visualization: Using vision to think*. Burlington: Morgan Kaufmann.
- Chen, M., Ebert, D., Hagen, H., Laramée, R., Van Liere, R., Ma, K.-L., et al. (2009). Data, information, and knowledge in visualization. *Computer Graphics and Applications*, 29(1), 12–19.
- Danis, C. M., Viegas, F. B., Wattenberg, M., & Kriss, J. (2008). Your place or mine: Visualization as a community component. In *Proceedings Human Factors in Computing Systems* (pp. 275–284).
- Hagen, P., Robertson, T., & Sadler, K. (2007). Accessing data: methods for understanding mobile technology use. *Australasian Journal of Information Systems*, 13(2), 135–149.
- Heer, J., & Agrawala, M. (2008). Design considerations for collaborative visual analytics. *Information Visualization*, 7(1), 49–62.

- Heer, J., Viegas, F. B., & Wattenberg, M. (2009). Voyagers and voyeurs: Supporting asynchronous collaborative visualization. *Communications ACM*, 52(1), 87–97.
- Isenberg, P., & Carpendale, S. (2007). Interactive tree comparison for co-located collaborative information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13(6), 1232–1239.
- Isenberg, P., Elmqvist, N., Scholtz, J., Cernea, D., Ma, K.-L., & Hagen, H. (2011). Collaborative visualization: definition, challenges, and research agenda. *Information Visualization*, 10(4), 310–326.
- Isenberg, P., Tang, A., & Carpendale, S. (2008). An exploratory study of visual information analysis. In *Proceedings of Human Factors in Computing Systems* (pp. 1217–1226).
- Kim, K., Javed, W., Williams, C., Elmqvist, N., & Irani, P. (2010). Hugin: A framework for awareness and coordination in mixed-presence collaborative information visualization. In *Proceedings of Interactive Tabletops and Surfaces* (pp. 231–240). New York: ACM.
- Ludwig, T., & Scholl, S. (2014). Participatory sensing im Rahmen empirischer Forschung. In *Mensch und Computer 2014—Tagungsband: Interaktiv unterwegs—Freiräume gestalten*, Hrsg. v. Butz, Andreas; Koch, Michael; Schlichter, Johann, München, Oldenbourg-Verlag, pp. 145–154.
- Mark, G., Kobsa, A., & Gonzalez, V. (2002). Do four eyes see better than two? Collaborative versus individual discovery in data visualization systems. In *Proceedings of Information visualization* (pp. 249–255).
- Viegas, F., Wattenberg, M., van Ham, F., Kriss, J., & McKeon, M. (2007). Manyeyes: A site for visualization at internet scale. *Transactions on Visualization and Computer Graphics*, 13(6), 1121–1128.
- Willett, W., Heer, J., Hellerstein, J., & Agrawala, M. (2011). Commentspace: Structured support for collaborative visual analysis. In *Proceedings of Human Factors in Computing Systems* (pp. 3131–3140).
- Wood, J., Wright, H., & Brodie, K. (1997). Collaborative visualization. In *Proceedings of Visualization* (p. 253–259).