A Mixed Approach for the Evaluation of a Guided Exploratory Visualization System

Nadia Boukhelifa[1] and Anastasia Bezerianos[2,1] and Evelyne Lutton[3]

[1]INRIA Saclay, France
[2]Univ Paris-Sud & CNRS (LRI), Saclay, France
[3]INRA, Grignon, France

Abstract

We summarise and reflect upon our experience in evaluating a guided exploratory visualization system. Our system guides users in their exploration of multidimensional datasets to pertinent views of their data, where the notion of pertinence is defined by automatic indicators, such as the amount of visual patterns in the view, and subjective user feedback obtained during their interaction with the tool. To evaluate this type of system, we argue for deploying a collection of validation methods that are: user-centered, observing the utility and effectiveness of the system for the end-user; and algorithm-centered, analysing the computational behaviour of the system. We report on observations and lessons learnt from working with expert users both for the design and the evaluation of our system.

Categories and Subject Descriptors (according to ACM CCS): H.5.2 [Information Interfaces and Presentation]: User Interfaces—User-centered design

1. Introduction

Exploratory visualization is a dynamic process of discovery that is relatively unpredictable due to the absence of a-priori knowledge of what the user is searching for [Gri96]. The focus in this case is on organising, testing, developing concepts, looking for trends, and defining hypothesis [Gri96]. When the search space is large, as is often the case for multidimensional data sets, the task of exploring and finding interesting patterns in data becomes tedious. Automatic dimension reduction techniques, such as PCA and MDS, reduce the search space, but are often difficult to understand, or require the specification of objective criteria to filter views before user exploration. Other techniques (e.g. [BLBC12]) guide the user to the most promising areas of the search space based on information learned during the exploration. This seems more adapted to the free nature of exploration.

In our previous work on guided exploratory visualization [CTBL12, TBBL13, BCTBL13], we tried to address the problem of helping users efficiently explore multidimensional datasets characterised by a large number of projections. We proposed a framework for Evolutionary Visual Exploration (EVE) that combines visual analytics with stochastic optimisation by means of an interactive evolutionary algorithm (Fig. 1). Our goal was to guide users to interesting projections, where the notion of "interesting" is defined by automatic indicators, such as the amount of visual patterns in a two-dimensional scatterplot, and by subjective user feedback obtained during their interaction with the system.
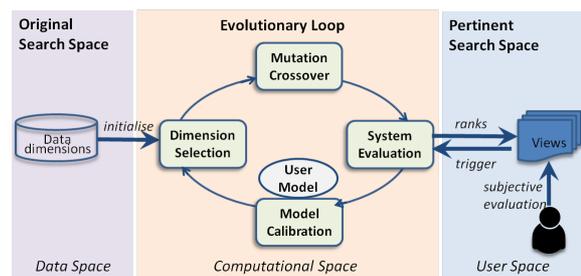
Figure 1: *The Evolutionary Visual Exploration framework: data dimensions are fed into an evolutionary loop in order to progressively evolve new interesting views to the user.*

2. On Evaluating Exploratory Visualization Systems

Similar to EVE systems, Dis-Function [BLBC12] allows users to interact with a visualization to guide machine learning and explore alternative structures in the data. Since their

focus was on algorithm performance, they evaluated their tool with non-expert participants using a simplified task and a generic dataset. Recently, Shao et al. [SBS*14] provided users with a sketch-based interface to input visual patterns of interest as they explored, but did not evaluate their system. User evaluations of exploratory systems, including guided ones, is challenging as exploratory tasks are fuzzy in nature, and expert users who can validate and perform these tasks are hard to recruit [Pla04, SMM12]. Moreover, case studies with experts are time consuming [Car08] and results may not be replicable and generalisable [Pla04]. Thus, rather than a single approach, Carpendale [Car08] recommends to take a variety of evaluative methodologies that together may start to approach the kind of answers sought. This indeed, was our intention when we designed case studies and experiments for our guided system as described in the next section.

3. Evaluation of an EVE System

To fully evaluate EVE, we felt that a collection of validation methods are needed, both user-centered, observing the utility and effectiveness of the system for end-users, and algorithm-centered, analysing the algorithmic behavior of the system. To this end we conducted (a) an observational study with five domain experts analysing their own data, and (b) a controlled study with twelve participants exploring synthetic data in order to examine in detail how users leverage the system and how the system evolves to match their needs.

3.1. User-Centered Evaluation

To assess the usability and utility of EVE, we tried to answer these three questions: (Q1) is our tool understandable and can it be learnt; (Q2) are experts able to confirm known insight in their data; and (Q3) are they able to discover new insight and generate new hypotheses. We designed three tasks: (T1) a game-task (similar to the task in section 3.2) with varying levels of difficulty to assess participants abilities to operate the tool; (T2) we asked participants to show in the tool what they already know about their data; and (T3) to explore their data in light of a hypothesis or research question they already had. This sequence of tasks assured that experts became familiar with the tool, and understood how to concretely leverage it by looking for known facts, before actually looking for new insights. Our evaluation approach sits between an observational study and an insight-based evaluation such as the one proposed by Saraiya et al. [SND05].

3.2. Algorithm-Centered Evaluation

We also conducted a controlled experiment to determine (Q4) whether the algorithm learns from user interactions and adapts to their change of focus. The task was designed as a game; a 5D dataset was synthesised with an embedded curvilinear relationship between two dimensions and noise for the rest of the dimensions. Participants were asked to quickly find a data projection that shows a derived visual pattern. We logged user interactions with the tool and the state of the system at each algorithm iteration. For the algorithmic analysis, we used both statistical and visualization techniques.

4. Discussion and Conclusions

We conducted qualitative and quantitative studies to evaluate EVE which helped us validate our framework of guided visual exploration. Our observational study led to interesting findings such as the ability of our tool to support experts in better formulating their research questions and building new hypotheses. For insight evaluation studies such as ours, reproducing the actual findings across subjects is not possible as each expert comes with their own dataset and questions. However, reproducing testing methodologies and coding for the analysis is. Although we run multiple field studies with experts from different domains, with sessions that were internally very different, the high level tasks, their order and the insight based coding were common. Training expert users on simple specific tasks that are not necessarily "theirs" also seemed to help experts become confident with the system, but of course comes at a time cost. While our quantitative study, allowed us to accurately describe the relationship between user behaviour and algorithm's response. Here the task was pre-defined and the data synthesised.

Guided visualization systems such as EVE fall under the wider arena of knowledge-assisted visualization [CH10] and mixed-initiative systems [Hor99]. In such cases, where the system is learning, it is crucial that users understand what the system is proposing or why changes are happening. Thus, when evaluating such systems with users, we need to specifically test if the automatic state changes and their provenance are understood. Research from the field of mixed-initiative systems describes a set of design principles that tries to address systematic problems with the use of automatic services within direct manipulation interfaces. These principles include considering uncertainty about a user's goal, transparency, and considering the status of users' attention [Hor99]. We can be inspired by the extensive experience and past work from HCI, to also consider how user behaviour can in turn adapt to fit our systems [MMPP00].

During the design, development and evaluation of EVE, we worked with domain experts at different levels. For the observational study, we worked with data experts from various disciplines which allowed us to asses the usefulness, usability and effectiveness of our system in different contexts. In particular, we largely benefited from having one domain-expert as part of the design and evaluation team. This expert explored multidimensional datasets as part of her daily work, using both algorithmic and visual tools. Involving end-users in the design team is a long-time tradition in the field of HCI as part of the user-centered design methodology. This is a recommendation we should consider more, both as a design and as a system validation approach. While HCI researchers acknowledge the challenges of forming partnerships with domain experts, their past experience (e.g. [CWK10]) can inform our community on how to proceed.

References

[BCTBL13] BOUKHELIFA N., CANCINO TICONA W. G., BEZERIANOS A., LUTTON E.: Evolutionary visual exploration: evaluation with expert users. *Computer Graphics Forum (EuroVis 2013, June 17–21, 2013, Leipzig, Germany) 32*, 3 (June 2013), 31–40. `doi:10.1111/cgf.12090`. 1

[BLBC12] BROWN E. T., LIU J., BRODLEY C. E., CHANG R.: Dis-function: learning distance functions interactively. In *IEEE VAST* (2012), IEEE Computer Society, pp. 83–92. 1

[Car08] CARPENDALE S.: Evaluating information visualizations. In *Information Visualization*, Kerren A., Stasko J., Fekete J.-D., North C., (Eds.), vol. 4950 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2008, pp. 19–45. URL: `http://dx.doi.org/10.1007/978-3-540-70956-5_2`, `doi:10.1007/978-3-540-70956-5_2`. 2

[CH10] CHEN M., HAGEN H.: Guest editors' introduction: knowledge-assisted visualization. *Computer Graphics and Applications, IEEE 30*, 1 (Jan 2010), 15–16. `doi:10.1109/MCG.2010.8`. 2

[CTBL12] CANCINO TICONA W. G., BOUKHELIFA N., LUTTON E.: EvoGraphDice: interactive evolution for visual analytics. In *Proceedings of the 2012 IEEE Congress on Evolutionary Computation* (Brisbane, Australia, June 2012), IEEE, (Ed.), Xiaodong Li, pp. 2286–2293. 1

[CWK10] CHILANA P. K., WOBBROCK J. O., KO A. J.: Understanding usability practices in complex domains. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2010), CHI '10, ACM, pp. 2337–2346. URL: `http://doi.acm.org/10.1145/1753326.1753678`, `doi:10.1145/1753326.1753678`. 2

[Gri96] GRINSTEIN G. G.: Harnessing the human in knowledge discovery. In *KDD* (1996), Simoudis E., Han J., Fayyad U. M., (Eds.), AAAI Press, pp. 384–385. 1

[Hor99] HORVITZ E.: Principles of mixed-initiative user interfaces. ACM Press, pp. 159–166. 2

[MMPP00] MACKAY W. E., MACKAY W. E., PROFESSOR V., PROFESSOR V.: Responding to cognitive overload: co-adaptation between users and technology. *Intellectica 30* (2000), 177–193. 2

[Pla04] PLAISANT C.: The challenge of information visualization evaluation. In *Proceedings of the Working Conference on Advanced Visual Interfaces* (New York, NY, USA, 2004), AVI '04, ACM, pp. 109–116. URL: `http://doi.acm.org/10.1145/989863.989880`, `doi:10.1145/989863.989880`. 2

[SBS*14] SHAO L., BEHRISCH M., SCHRECK T., VON LANDESBERGER T., SCHERER M., BREMM S., KEIM D. A.: Guided sketching for visual search and exploration in large scatter plot spaces. In *Proc. EuroVA International Workshop on Visual Analytics* (2014). 2

[SMM12] SEDLMAIR M., MEYER M., MUNZNER T.: Design study methodology: reflections from the trenches and the stacks. *IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis) 18*, 12 (2012), 2431–2440. 2

[SND05] SARAIYA P., NORTH C., DUCA K.: An insight-based methodology for evaluating bioinformatics visualizations. *IEEE Transactions on Visualization and Computer Graphics 11*, 4 (July 2005), 443–456. URL: `http://dx.doi.org/10.1109/TVCG.2005.53`, `doi:10.1109/TVCG.2005.53`. 2

[TBBL13] TICONA W. C., BOUKHELIFA N., BEZERIANOS A., LUTTON E.: Evolutionary visual exploration: experimental analysis of algorithm behaviour. In *GECCO (Companion)* (2013), Blum C., Alba E., (Eds.), ACM, pp. 1373–1380. 1