An Experiment and Analysis System Framework for the Evaluation of Contextual Relationships

Ralf Bierig
Rutgers University
bierig@rci.rutgers.edu

Michael Cole
Rutgers University
m.cole@rutgers.edu

Jacek Gwizdka
Rutgers University
cirse10@gwizdka.com

Nicholas J. Belkin
Rutgers University
belkin@rutgers.edu

Jingjing Liu
Rutgers University
jingjing@eden.rutgers.edu

Chang Liu
Rutgers University
imliuc@gmail.com

Jun Zhang
Rutgers University
zhangj@eden.rutgers.edu

Xiangmin Zhang
Rutgers University
xiangminz@gmail.com

ABSTRACT
This paper presents an experiment and analysis system framework that allows researchers to design and conduct interactive experiments and analyze data for the evaluation of contextual relationships.

Categories and Subject Descriptors
H.4 [**Information Systems Applications**]: Miscellaneous

Keywords
Context Evaluation, Information System, Data Analysis

1. INTRODUCTION AND BACKGROUND
In the last decade, context-aware computing has made much effort to formalize context[3], describe general context models[7] and develop systems that apply such models in different application domains [5] – such as mobile computing (e.g. tourism and recreation [16, 11]). There is, however, only limited research about the experimental evaluation of context, particularly about the effects of various contextual attributes and their interaction. This gap is beginning to be addressed with several workshops and conferences [9, 8, 12, 2].

Rigorous experimentation in this domain presents challenges in that such experiments are generally difficult to administer and demanding in resources [6, 10]. Although software frameworks for contextual enrichment of applications exist [13, 4] there is generally little system-related support for comprehensive evaluation of context attributes and models. This paper presents a system framework that provides researchers with a tool to: 1) *design and conduct experiments* for the evaluation of particular contextual attributes and 2) *integrate data and analyze results* to better understand contextual relationships. The framework promotes an interactive and task-oriented viewpoint that is supported by a wide range of logging tools.

Section 2 reviews the system architecture consisting of the experiment and the analysis system. Section 3 describes how the architecture supports researchers to investigate and evaluate contextual relationships. Section 4 discusses the current state of the system and future plans for its dissemination.

Appears in the Proceedings of The 2nd International Workshop on Contextual Information Access, Seeking and Retrieval Evaluation (CIRSE 2010), March 28, 2010, Milton Keynes, UK.
http://www.irit.fr/CIRSE/

2. SYSTEM FRAMEWORK
The system framework is part of a project deliverable[1] that aims to investigate ways to improve users' ability to find information in search environments such as digital libraries. In particular we analyze various interacting contextual factors that are involved in such online search activities. Despite our focus, results are expected to contribute to a much wider range of application environments such as mobile search and recommender systems.

2.1 Overview
The overall aim of our framework is to reduce the complexity of designing and conducting experiments and integrating and analysing results from experiments for the evaluation of contextual relationships as usually expressed in user and context models. Such experiments usually require a complex arrangement of system components (e.g. GUI, user management and persistent data storage). Our framework enables researchers to focus on research related issues (e.g. task and questionnaire design and the selection of experiment variables) rather than the creation of the experiment logic and the transformation, integration and the processing of data and results after the experiment has been completed. This helps to reduce the overall time and effort that is needed to design and conduct experiments and to get valuable results about contextual relationships from experiment data. As shown in figure 1, the system framework consists of two parts

[1] http://comminfo.rutgers.edu/imls/poodle/

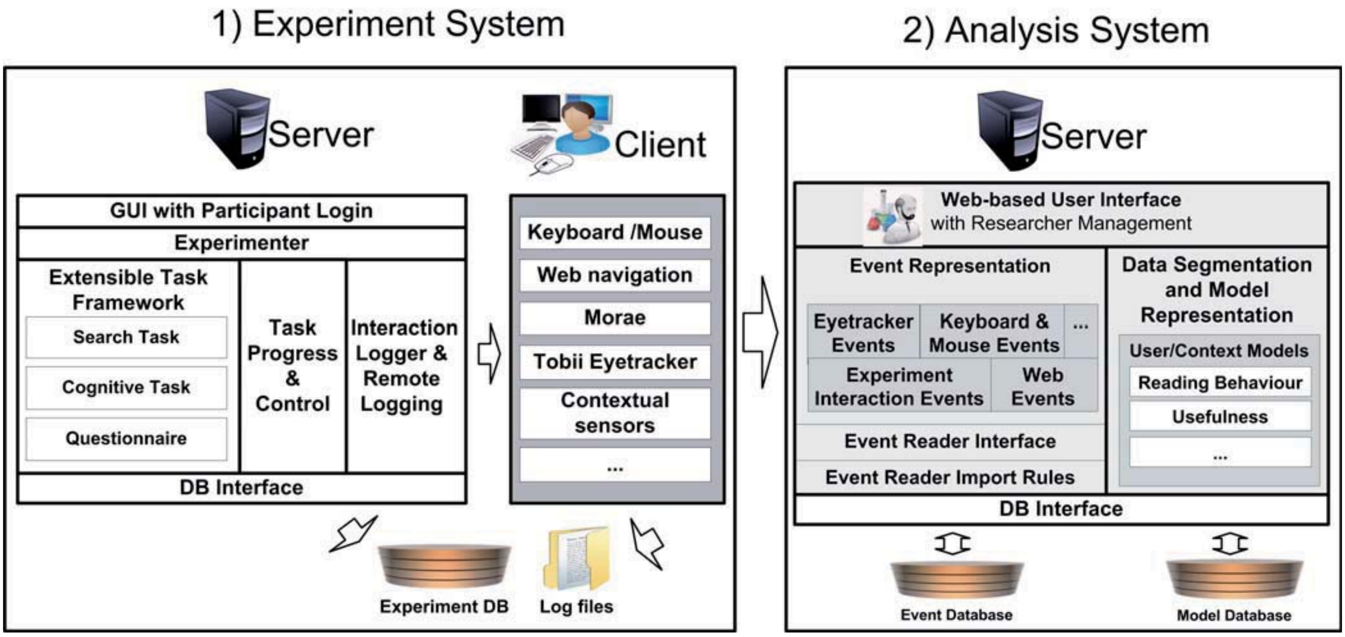

Figure 1: Components of the experiment and analysis system framework

– 1) an *experiment system* that allows researchers to design and conduct interactive experiments in close-to-operational application environments and 2) an *analysis system* that enables them to integrate and analyze results obtained from such experiments.

2.1.1 Experiment System

The experiment system, described in more detail in [1], includes a number of components.

The *GUI* provides authenticated login for participants, their assignment to one or more experiments and basic navigational support during an experiment. The *Experimenter* controls and coordinates an *Extensible Task Framework* that offers researchers a set of reusable tasks that can be used for creating experiments (e.g. standard open web search tasks). Own tasks can also be added to this collection. Furthermore, the Experimenter manages *Task Progress and Control* that balances task sequences, monitors the progress of participants including the safe recovery of interrupted sessions. In addition, the *Interaction Logger with Remote Logging* provides a mechanism for tasks to log contextual information internally at specific points during an experiment task and to call external logging applications on the client. This allows creating more effective experiments that may include different kinds of contextual data logging on both server and the client side. Whereas the server has a central logging facility, the client consists of a flexible and expandable array of independent loggers. Currently, these loggers observe the most commonly known user behaviours – keyboard and mouse activities, web navigation, usability information from Morae[2] and eye-tracking data from Tobii[3]. This list can easily be expanded with other (existing or new) logging tools that cover additional contextual information from the user or the user's environment. Examples may include location information (e.g. geographic position or proximity to points of interest) or physiological states of the user (e.g. heart rate or Galvanic skin response). Logging information is either stored in an experiment database through the *DB Interface* or in application-specific log files.

2.1.2 Analysis System

The analysis system serves as an extension of the experiment system with additional features to integrate experiment data into a unified data structure. Researchers can inspect and explore these data sets and segment and model results to gain a better understanding of contextual relationships. The analysis system consists of the following components:

- The *Event Representation* integrates experiment data through the *Event Reader Interface* into a unified event data structure. This data structure is extensible and the collection of event readers mirror the logging tools provided with the experiment system as described in the previous section. An extensible set of event types ensures that researchers can adapt and extend the analysis framework to process data from a variety of experiments under a single platform. This ensures that additional logging tools can be introduced through the experiment system to capture additional types of user context either through the logging of high-level user behaviour or through the application of low-level sensors as described in [14].
- *Event Reader Import Rules* can be used to configure event readers and therefore adapt the data import process. Such rules can for example be applied to add additional filters for event readers (e.g. excluding web

[2] http://www.techsmith.com
[3] http://www.tobii.com

events with certain URLs) or providing standard validation (e.g. tagging certain events as problematic thus flagging results for manual inspection).

- *Data Segmentation* divides experiment data into semantic units guided by research hypotheses. The system framework provides a standard minimal segmentation by distinguishing data based on experiments, users and tasks. The research can add additional levels of data segmentation to structure data in smaller logical units. A segmentation can for example differentiate interaction data based on users' current stage in a search task (e.g. distinguishing users' task stages of query formulation, result page inspection and content page viewing) or, more generally, data can be segmented along low-level decision points (e.g. mouse clicks and/or key strokes).

- *Model Representation* processes (segmented) event sequences to test specific research hypotheses i.e. verifying effects of context attributes and relationships between them (e.g. identifying users' perceived usefulness of content and determining reading behaviour). Other data segmentations and model representations can be added by researchers to further specialize the system framework for particular types of analysis.

- The *Web-based User Interface* extends the system to an online service where researchers can generate, inspect and share event representations, data segmentations and models within one or across multiple experiment data sets. These are stored through a *DB Interface* that persists both event and model representations into separate databases for later reuse. The user interface supports authenticated login to allow the system to be used as part of a collaborative research platform.

3. CONTEXT EVALUATION WITH THE SYSTEM FRAMEWORK - BENEFITS AND LIMITATIONS

The system design incorporates many aspects useful for the evaluation of contextual relationships from data obtained in interactive and task-based experiments. This section summarizes these aspects, shows how they relate to the system framework, points out how they can help researchers to evaluate context, and expresses limitations that should be considered.

- *Modularity:* Context models may cover a wide range of attributes based on dimensions such as the application environment (e.g. library or mobile environment) and the intended user group (e.g. professional journalists or online web searchers) as well as others. The system framework supports this requirement in a number of ways. First, a modular and multi-dimensional logging framework within the experiment system can record behavioural data from the user and sensory data from the user's environment. Second, these multi-dimensional data streams can be integrated into a unified stream of events within the analysis system. Third, this event stream can be treated holistically through segmentation, as a tool for data categorization and conditioning, and through modelling to investigate and discover contextual relationships.

- *Extensibility:* As an extensible framework with respect to contextual logging tools (in the experiment system) and readers, rules, segmentations and models (in the analysis system) the framework offers researchers ways to adapt and extend it to their own requirements and research agendas. These extensions however require additional, customizing implementation work by the user of the system framework; for example adding another logging tool to measure a new contextual aspect from the user also requires implementing the corresponding event representation and an additional reader to import the new data log. Such procedures, however, are guided through the application of programming interfaces and supported with examples that are available in open source as part of the project. This is not much different from other extensible software frameworks such as WEKA [15].

- *Separation between data and modelling:* Data (in the form of low-level event representations) is separated from its interpretation (in the form of high-level segmentations and models). Thus, it is possible to generate multiple, alternative context models from the same underlying events that can each be evaluated in isolation. This also allows user and context models to be reused for different data segments from one or across multiple experiments.

- *Collaboration* is central to the design and has been supported in both parts of the system framework. The experiment system allows researchers to implement and share experiment tasks thus building a collaborative repository (e.g. internet search tasks, tag cloud search, standard questionnaires for language understanding and various cognitive tests). Likewise, configurations for behavioural and contextual logging tools can be created and reused across different experiments and shared between researchers. The analysis system offers a meeting platform through its web-based user interface. Data, segmentations and models can be configured, integrated and shared between researchers allowing collaborating with data and ideas and forming virtual research groups. Researchers can create and exchange integrated event data sets from experiments specific to the needs of individuals or groups (e.g. event data limited to a subset of experiment participants, experiment tasks or types of context such as web activity or eye movement). Shared data sets can then be applied for further data segmentation (e.g. selecting only particular user activities or contextual states, such as query input or reading behaviour). An extensible pool of models can be applied to such segments and accessed collaboratively. Basic summary visualizations are available and findings can be exported allowing researchers to further process data with third-party tools and apply results (e.g. integrating a learned context model in a personalized desktop search application).

4. CURRENT STATE AND FUTURE PLANS

A prototype of the experiment system has has been designed and developed with active work on improving logging comprehensiveness (especially for contextual, sensor-based logging) and scalability. The experiment system has already been applied to design and conduct four experiments each with distinctive design and goals for our research project. In those experiments we have collected rich contextual information for the basic investigation of relationships between use behaviour and various user context attributes such as cognitive abilities and individual differences, reading and scanning behaviour and perception of usefulness during online search. The analysis system has been designed and the modelling and user interface is in active development. The experiment system framework has been released as open source[4]. The analysis system will be released as open source when it is feature complete and stable. Both of these systems can benefit the research community by allowing for collaboration between researchers and enabling additional improvements and extensions to better serve the needs of context researchers.

Acknowledgements: This work is supported by IMLS grant LM-06-07-0105-07.

5. REFERENCES

[1] R. Bierig, J. Gwizdka, and M. Cole. A user-centered experiment and logging framework for interactive information retrieval. In N. J. Belkin, R. Bierig, G. Buscher, L. v. Elst, J. Gwizdka, J. Jose, and J. Teevan, editors, *SIGIR 2009 Workshop on Understanding the user - Logging and interpreting user interactions in IR*, Boston, MA, 2009.

[2] P. Borlund, J. W. Schneider, M. Lalmas, A. Tombros, J. Feather, D. Kelly, A. de Vries, and L. Azzopardi. Second symposium on information interaction in context (iiix). London, UK, 2008. ACM Press.

[3] A. K. Dey, G. Kortuem, D. Morse, and A. Schmidt. Special issue on situatuated interaction and context-aware computing. *Personal and Ubiquitous Computing*, 5(1), 2001.

[4] P. Fahy and S. Clarke. Cass - middleware for mobile context-aware applications. In *2nd International Conference on Mobile Systems, Applications, and Services (MobiSys 2004), Workshop on Context-Awareness*, Bosten, MA, USA, 2004.

[5] A. Göker, H. I. Myrhaug, and R. Bierig. Context and information retrieval. In A. Göker and J. Davies, editors, *Information Retrieval: Searching in the 21st Century*. John Wiley and Sons, Ltd, Chichester, UK, 2009.

[6] J. Goodman, S. Brewster, and P. Gray. Using field experiments to evaluate mobile guides. In *6th International Symposium on Human Computer Interaction with Mobile Devices and Services (Mobile HCI), International Workshop on HCI with Mobile Guides*, Glasgow, UK, 2004.

[7] J. Indulska and D. D. Roure. Workshop on advanced context modelling, reasoning and management. In *6th International Conference on Ubiquitous Computing (UbiComp)*, Nottingham, UK, 2004.

[4] `http://sourceforge.net/projects/piirexs/`

[8] P. Ingwersen, K. Jaervelin, and N. Belkin. Workshop on information retrieval in context. In *28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Salvador, Brazil, 2005. Royal School of Library and Information Science, Copenhagen, Denmark.

[9] P. Ingwersen, K. van Rijsbergen, and N. Belkin. Workshop on information retrieval in context (irix). In *27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Sheffield, UK, 2004.

[10] J. Kjeldskov, C. Graham, S. Pedell, F. Vetere, S. Howard, S. Balbo, and J. Davies. Evaluating the useability of a mobile guide: The influence of location, participants and resources. *Behaviour and Information Technology*, 24(1):51–65, 2005.

[11] D. M. Mountain and A. MacFarlane. Geographic information retrieval in a mobile environment: Evaluating the needs of mobile individuals. *Journal of Information Science*, 33(5):515–530, 2007.

[12] I. Ruthven, P. Borlund, P. Ingwersen, N. Belkin, A. Tombros, and P. Vakkari, editors. *Information Interaction in Context. 1st International Symposium on Information Interaction in Context, IIiX 2006*. ACM Press, Copenhagen, Denmark, 2006.

[13] D. Salber, A. K. Dey, and G. D. Adowd. The context toolkit: Aiding the development of context-enabled applications. In *Conference on Human Factors in Computing Systems (CHI)*, pages 434–441, Pittsburgh, PA, USA, 1999. ACM Press.

[14] A. Schmidt. *Ubiquitous Computing - Computing in Context*. Phd thesis, Lancaster University, 2002.

[15] I. H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 2nd edition, 2005.

[16] A. Zipf. User-adaptive maps for location-based services (lbs) for tourism. In *9th Int. Conf. for Information and Communication Technologies in Tourism (ENTER 2002)*, pages 329–337, Innsbruck, Austria, 2002. Springer Verlag.