AI Methods for Smart Environments

A Case Study on Team Assistance in Smart Meeting Rooms

Martin Giersich, Thomas Heider, Thomas Kirste

Department of Computer Science, Rostock University
Albert-Einstein-Straße 21
18059 Rostock, Germany
first.last@uni-rostock.de

Abstract. Ubiquitous computing aims for the realisation of environments that assist users autonomously and proactively in a non-distractive manner. Therefore smart environment infrastructures need to be able to identify users needs (*intention recognition*) and to plan an appropriate assisting strategy (*strategy generation*) without explicit user interaction. In our two-stage approach we address inferring the intention of a team of users during a meeting within a smart multiple display environment and the system decision process – what information to present on which display – on the strategy generation level.

1 Introduction

A central requirement for an assistance architecture for a smart meeting room is that it should support technical infrastructures that are built from individual components in an ad hoc fashion. Our solution approach is a two-stage design, where at the first stage the system components recognize the intention of a team of users, and at the second stage, the system components jointly generate a strategy that fulfills the needs of the team. In this paper we report the results of our current research and the ongoing evaluation. Part one of the paper presents results in intention analysis and part two represents strategy generation and evaluation. At this time we evaluated the parts separately. An evaluation of the complete integrated system will be matter of future work.

Intention recognition becomes a challenge, especially if multiple users are observed by noisy heterogenous sensors. We propose a *team behavior model* based on hierarchical dynamic Bayesian network (DBN) for inferring the current task and activity of a *team* of users. Given (noisy and intermittent) sensor readings of the team members' positions in a meeting room, we are interested in inferring the team's current objective.

A simulation data evaluation of our particle filter based team behavior model shows reasonable inference accuracy and speed for our implemention and demonstrates how additional unreliable knowledge about the meeting agenda improves prediction accuracy and speed. Here, we claim that even unreliable agendas improve intention recognition in smart environments for a compliant team behavior without sacrificing recognition accuracy for the non-compliant case.

We propose to cast the **Strategy generation** problem as an optimization task. As example problem we use the document-display mapping question, which is what to present on what display in a multi-user, multi-display environment. We suggest the definition of an *explicit global quality measure* to achieve coherent ensemble behavior for a team of multiple users with (maybe) diverging interests.

The evaluation of this part shows, that an automated document display mapping based on an explicit global quality measure leads to coherent ensemble behavior and is at least as effective as conventional manual assignment, while at the same time significantly reducing the number of required interactions. This claims are based on user performance data collected in the scope of a comparsion study.

2 Intention Recognition

Especially for the intention recognition used in our prototype smart meeting room we studied whether incomplete and unreliable (i.e., sometimes misleading) knowledge about the needs of a team of users (agenda) can be used to improve the quality of intention recognition. Specifically, we were interested in the usefulness of an unreliable agenda for improving the recognition of team activities during a meeting. Based on Bayesian filtering and an explicit probabilistic team behavior model we have carried on a simulation study that allowed us to answer the following questions:

– How accurate and how fast can we predict team behavior with an agenda assumption and history knowledge?
– What influence do deviations of the team from the planned agenda assumption have on prediction quality (i.e., does a wrong agenda degrade the quality of intention recognition)?
– How flexible does an agenda assumption need to be in order to optimally predict team behavior?

We chose simulation of data rather than real world data as this enabled us to configure the probability distribution of the sensor readings. We used Gaussian and Cauchy distributed sensor readings with a variety of different parameter settings to examine the influence of the sensor model on the prediction quality.

Team Behavior Model – Bayesian Filtering for identifying a user's current task has been successfully used in several projects that aimed at supporting user activities in classrooms, meeting rooms, and office environments [1–3]. Here, *dynamic Bayesian networks* (DBNs) were investigated increasingly for modeling a user's activities [4, 5]. In our own work, we looked at using DBNs for inferring the current task and actions of a *team* of users. Given (noisy and intermittent)

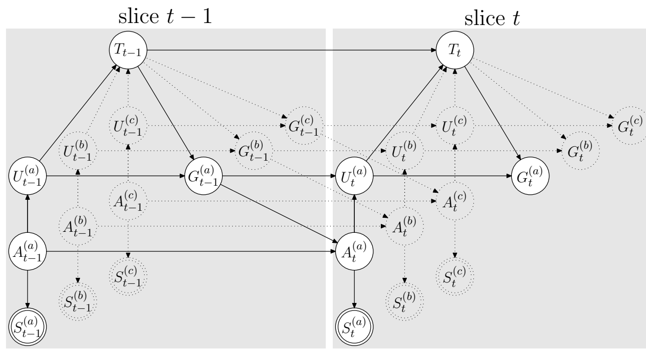

Fig. 1. Two-sliced dynamic Bayesian network (DBN) modeling team intention inference. It shows the intra-slice dependencies between observable (double-contoured) and hidden variables, as well as the inter-slice dependencies between consecutive states.

sensor readings of the team members' positions in a meeting room, we were interested in inferring the team's current objective – such as having a presentation delivered by a specific team member, a moderated brainstorming, a round table discussion, a break, or the end of the meeting.

The basic structure of the DBN we propose for modeling the activities of such a team is given in Figure 1. With this DBN we try to model the behavior of a team of three users during a meeting. In order to exploit agenda information, we need a DBN structure that is able to incorporate an explicit agenda, and that represents the negotiation process between the team and its members during activity selection. At the top level, the team node T_t represents the current team intention. The team's intention at time t depends on what the team has already achieved (T at time $t-1$, T_{t-1}), and what the users i are currently trying to achieve (the $U_t^{(i)}$-nodes, $i \in \{a, b, c\}$). The $G_t^{(i)}$ nodes represent the new individual assignments if the team T will adopt a new intention. So at each time slice, the team looks at what the users have achieved so far and then decides what the users should do next. What the user is doing at time t depends on his previous action (e.g., the user's current position and velocity) and assignment – $A_{t-1}^{(i)}$ and $G_{t-1}^{(i)}$. Finally, the sensor observations of user i at time t – the nodes $S_t^{(i)}$ – depend on the user's activities at that time.

Note that these sensor nodes are the *only* observable nodes in our model: we estimate the team's negotiations from the observable behavior of the team members. Once a probabilistic model is available, it allows us to infer user and team intentions.

Experimental Design and Results – Clearly, agenda information should improve the quality of team intention recognition. However, as soon as a team deviates from the a-priori agenda, recognition quality may drop: The recognizer may be led to wrong conclusions by misleading a-priori information that potentially defeat any benefit. Objective of our evaluation has been to investigate,

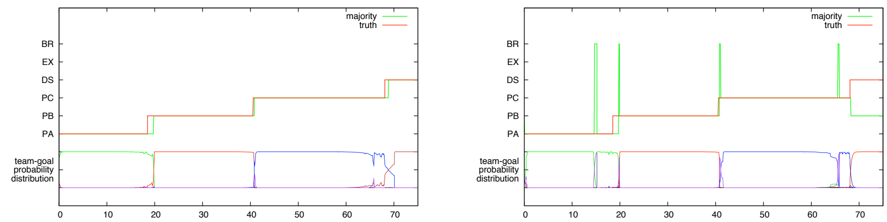

Fig. 2. Inference of a $\langle \mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D} \rangle$ truth from Cauchy distributed sensor data (delay 0.25, error 10.0) with the trackers $T_{.8}$ (left) and $T_{uniform}$ (right).

whether a-priori agenda information can be used to improve recognition quality in case the team complies to the agenda, *without* sacrificing recognition quality in case of non-compliance with the agenda. We were interested in two main questions:

(a) How reliable is agenda based recognition in case of compliance and non-compliance, compared to an agenda-less tracking?
(b) How fast will an agenda based recognizer identify a change in the team objective for these cases?

To analyze the effect of an agenda on reliability and speed of intention recognition in case of compliance and non-compliance we chose three different conference sequences (one compliant, two non-compliant). Further we used four different parameter settings for the sensors. In two settings we used sensor data that is Gaussian distributed. The two other settings sensor data followed a Cauchy distribution. The settings for each distribution differed in delay between consecutive sensor readings and sensor error.

For the evaluation of recognition accuracy, we used four different models for a-priori agenda information – a random model where every activity has the same probability and history is not tracked ($T_{uniform}$) and three models with different start probabilities for user A $\{.6, .8, .95\}$ and the other users respectively ($T_{.6}$, $T_{.8}$, $T_{.95}$). For every tracker model six runs were logged. The illustration of two typical representative of model $T_{.8}$ and model $T_{uniform}$ simulation runs in Figure 2 shows that the main uncertainty about the teams objective pervails during the phase of an objective shift. The left picture shows the advantage of agenda knowledge. For instance the objective shift from *B Presents* (PB) to *C Presents* (PC) around time slice 40 is recognized faster and more reliable. Further it shows that agenda knowledge leads to less misinterpretation of sensor readings. So the overall error rate shrinks. Figure 3 shows solid recognition also for non-compliant cases. Here, tracked with model $T_{.8}$.

The averages over 6 simulation runs for 48 different parameter settings give an delay between true objective shift of the team and the recognition of this shift of $7.36sec$ for $T_{.8}$ versus $10.95sec$ for $T_{uniform}$. The average intention recognition reliability for the best model $T_{.8}$ was measured with 91.16% correct versus 83.1%

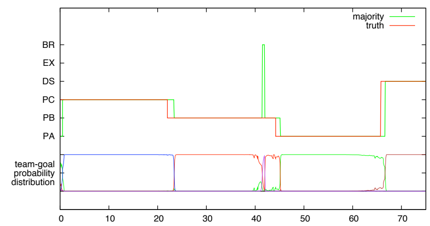

Fig. 3. Inference of the non-compliant truth $\langle \mathsf{C}, \mathsf{B}, \mathsf{A}, \mathsf{D} \rangle$ from Cauchy distributed sensor data (delay 0.25, error 10.0) with the trackers $T_{.8}$.

for the uniform model. Comparison of the reliability values for $T_{.8}$ and $T_{uniform}$ gives the most important result of this study:

> It is possible to improve the recognition accuracy for the compliant case by using an agenda, *without* sacrificing recognition accuracy for the non-compliant case.

Therefore, it always pays to include available a-priori agenda information in the recognition system, even if the correlation between the agenda sequence and the true activity sequence is not very strong.

However, it is important to assign a suitable probability to the agenda's preferred sequence. If this value is too high (e.g., .95), the agenda becomes too rigid: it will tend to assume that the team follows the agenda, even if the sensor data does tell a different story. On the other hand, further increasing the looseness of the agenda (e.g., to .6) does not improve the recognition of the non-compliant action sequences. We suspect that unnecessary looseness will eventually degrade recognition capability, but we have not observed this in our data.

Finally, simulation results show that an agenda reduces the delay, specifically for the later team actions. (Clearly, the agenda will not reconsider items already worked off, an aspect favorably reducing the degrees of freedom in comparison to $T_{uniform}$.)

3 Strategy Generation

Multi-display environments support collaborative problem solving and teamwork by providing multiple display surfaces for presenting information [6, 7]. One difficulty here is the *display mapping problem* – that is, deciding which information to present on what display in order to optimally satisfy the users' needs for information. Current approaches for controlling multi-display environments rely on manual assignment [8, 9], using a suitable interactive interface and resolving conflicts by social protocols (negotiations). However, manual display assignment has to cope with the following problems:

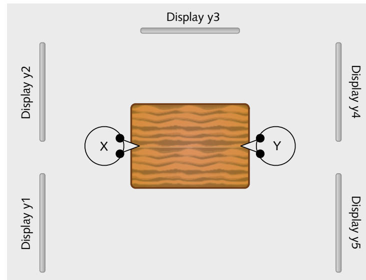

Fig. 4. Experimental setup for strategy generation comparsion study

– **Interest conflicts** between users might be solved faster by computer supported negotiation mechanism: It was observed that social protocols do not always suffice for coordinating the use of shared resources [10].
– The need for **dynamic realignment** of display mapping is caused by topic changes in the user population: In this situation, the user's focus of attention will be on the changing topic rather than on convincing the display infrastructure to change the topic.

So, an automatic display assignment might be helpful in multiple display environments, specifically in multi-user settings. However, to our knowledge, it is not known if suitable automatic assignment heuristics can be found. This is the question we want to answer.

Display Mapping Quality Measure – A *display mapping* is a function m, which assigns documents to sets of displays. For a given document d, $m(d)$ gives the set of displays document d is assigned to. In order for automatic display mapping to be successful it is necessary to identify a well-defined quality measure that sufficiently captures the users needs. Clearly, at least the following aspects are reasonable:

Spatial Layout: For documents of high importance to a user, displays should be preferred that provide a good visibility for the user. Formally, this critierion for m can be defined as

$$q_s(m) = \sum_{\substack{u \in U \\ d \in D}} impt(d,u) * \max_{y \in m(d)} vis(y,u) \tag{1}$$

where $impt(d,u) \in [0\,..\,1]$ denotes the importance of the document d to a user u, and $vis(y,u) \in [0\,..\,1]$ the *visibility* of display y by user u. If a document is assigned to multiple displays, only the best one ("primary display") for a given user is considered when computing the quality for this user (this is the "max vis" term). Note, that deriving a reliable estimation of *impt* in general may be a substantial challenge. We think that additional informations available from intention recognition can be used as a surrogate (such as agenda listings, team members roles and associated documents, etc.).

Temporal Continuity: When considering a display for a document, the system should prefer already existing assignments: Documents should not unnecessarily change their place. A relevant display shift occurs between two mappings, if a user's primary display for a document changes. We then try to minimize these shifts relative to the document's importance. Based on these criteria, we have developed an algorithm that is able to automatically compute a display mapping for a set of users and documents (see [11]).

Experimental Design – The objective of our evaluation experiment was to measured the impact of manual vs. automatic display assignment on the performance of a team in solving a semi-cooperative task. In such tasks, the need of cooperation and joint use of information is not evident from the start, but rather arises while working on the task. We think that this kind of aspect pertains to many team processes.

Two-person teams had to solve a semi-cooperative set of comparison tasks as fast as possible. The two team members, X and Y, were given different agendas, each containing the description of an individual comparison. For X the task was to do a simple letter comparison of two documents `A` and `B`, for Y the task was to compare `A` and `C`. In addition, X and Y had to report time information and a random key from another document `Time`. The seemingly unrelated tasks for X and Y were linked into a cooperative task through the shared documents `A` and `Time`.

Every participant was given a simple user interface for document assignment. Manually assignment of a document to a display-surface is done through simple "drag & drop". For automatic assignment, the user just associates an importance value with the documents. As the agendas and task descriptions were mutually unknown, the sharing had to be discovered through a conflict in the manual assignment group.

For each experiment, we recorded the time required for completing the task, the number of interactions and the solution correctness (percentage of letter differences found). After each task set, the subjects were asked to answer a questionnaire regarding user satisfaction. After both task sets, the subjects were asked to complete a final questionnaire regarding the comparison of automatic versus manual assignment.

24 voluntary subjects were recruited from staff members and students of the local university. The teams had to solve two sets comparison tasks in sequence. Group A had to solve the first set using automatic assignment and the second set with manual assignment. The Group M was given the tasks in reverse order. In the evaluation of the results, we will call the first set "Initial Test" and the second "After Training", respectively. (See [12] for a more detailed discussion of both experimental setup and findings.)

Findings – When the teams were using automatic assignment, the average time to complete one set of a comparison task was 4:08*min*, while they required an average time of 4:49*min* using manual assignment. The subjects needed 8.5 interactions on average with automatic and 15 interactions on average with manual

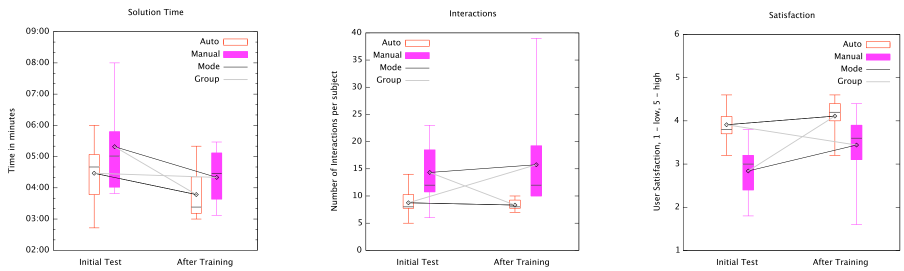

Fig. 5. Boxplots of solution time vs. mode, overall (left); interaction count vs. mode (middle); user satisfaction vs. mode (right)

assignment. This indicates that the automatic assignment is superior to manual assignment, regarding time and interactions.

An overview of the collected data is shown in the boxplots[1] in Figure 5. In these plots, "Mode" refers to the display assignment mode (manual vs. automatic). In the per-task-set plots, grey lines connect the mean values of the two consecutive task sets of a group (Group A or Group M), black lines connect consecutive task sets using the same assignment mode.

As can be seen in Figure 5, left, for both task sets the solution time is shorter when using automatic assignment. In addition, Group M was able to solve the task substantially faster in the second set (i.e., when switching from manual to automatic assignment), whereby Group A was not able to improve performance in the second set (i.e., switching from automatic to manual assignment). The number of interactions (Figure 5, middle) is smaller for the automatic method in both sets.

In the manual assignment mode, both groups initially had no idea that they needed to share documents. So they unwittingly "stole" the shared documents from each others "private" displays. It took a couple of interactions until the participants realized that they needed to cooperate and to assign some of the documents to a display visible to both users. This process of realization and negotiation was the reason for confusion and delay.

In the automatic assignment mode no such conflicts did arise as the system automatically displayed shared documents on a shared screen. If we use the number of interactions as indicator of occurred conflicts, the data shows that with the automatic mode the number of conflicts is considerably smaller than in the manual mode. A detailed survey of the log files showed that documents which had to be shared, very frequently were reassigned in the manual mode. This proves the presumption that resolving conflicts by social negotiation is – in some situations – inferior to a computer supported negotiation, which can be solved by an automatic assignment using a global quality function such as q.

[1] These boxplots show the minimum and maximum values, the 25% and 75% percentiles, the median (horizontal bar inside the box), and the mean (small circle inside the box).

For assessing user satisfaction, we used parts of the technology acceptance model (TAM)[13]. We included the following items, each to be answered on a scale from 1 (strongly disagree) to 5 (strongly agree):

> *The system is easy to use. – The system helps in solving the task efficiently. – It is easy to cooperate with the team partner. – The system helps in solving team conflicts. – I felt comfortable in using the system.*

The final questionnaire used the same items with the request to compare both approaches, automatic and manual assignment, on a scale from 1 (manual assignment strongly preferred) to 5 (automatic assignment strongly preferred).

The distribution of the user satisfaction data using per-questionaire averages is shown in Figure 5 (right). The overall user satisfaction is higher in the auto mode, for both task sets. In addition, user satisfaction *decreases* within a *group* when switching from auto to manual, while it *increases* when switching from manual to auto.

The correlation of the subjective user satisfaction with the objective data from the log files confirm our hypothesis that the automatic display assignment is superior to the manual assignment in multi-user, multi-display situations with conflicting and dynamic document sets.

4 Summary

We have discussed the problem of assisting teams in effectively using multi-display environments for working together and we have addressed the question whether it is possible to infer the intention of the team and to find well-defined quality criteria for automatic display assignment.

Our results regarding team intention recognition, inference accuracy and speed showed that despite noisy observable sensor data and a rather ad hoc prior probability distribution for the occurrence of agenda items a precise and robust inference is possible. Further adding agenda knowledge to a team behavior model was identified as improvement for the compliant case and as non-disturbing for the non-compliant case. So, we can claim that unreliable agendas are useful for inferring team intentions. We will now focus on in-depth development of an appropriate team behavior model and incorporate learning of probabiltiy distributions using *EM-algorithm.*

In the strategy generation part, we have been able to show that automatic assignment enables teams to solve their tasks in a shorter time, with less conflicts between team members, with greater satisfaction and with reduced cognitive load. Future investigations will have to show whether this benefit offers the universality and significance required to incorporate it generally into smart multiple display environments.

Finally the seamless integration of our two-stage design is an issue that we will address in the future work.

References

1. David Franklin, Jay Budzik, and Kristian Hammond. Plan-based interfaces: keeping track of user tasks and acting to cooperate. In *Proceedings of the 7th international conference on Intelligent user interfaces*, pages 79–86, New York, NY, USA, 2002. ACM Press.
2. Hung Bui. A general model for online probabilistic plan recognition. In *IJCAI '03: Proceedings of the 18th International Joint Conference on Artificial Intelligence*, pages 1309–1315, 2003.
3. Thi V. Duong, Hung H. Bui, Dinh Q. Phung, and Svetha Venkatesh. Activity recognition and abnormality detection with the switching hidden semi-markov model. In *CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 1*, pages 838–845, Washington, DC, USA, 2005. IEEE Computer Society.
4. Donald J. Patterson, Lin Liao, Dieter Fox, and Henry A. Kautz. Inferring high-level behavior from low-level sensors. In *Proceedings of UBICOMP 2003: The Fifth International Conference on Ubiquitous Computing*, October 12-15 2003.
5. Donald J. Patterson, Lin Liao, Krzysztof Gajos, Michael Collier, Nik Livic, Katherine Olson, Shiaokai Wang, Dieter Fox, and Henry A. Kautz. Opportunity knocks: A system to provide cognitive assistance with transportation services. In *Ubicomp*, pages 433–450. Springer, 2004.
6. Alois Ferscha, Gerd Kortuem, and Antonio Krüger. Workshop on ubiquitous display environments. In *Proc. Ubicomp 2004*, Nottingham, England, Sep 7 2004.
7. Maribeth Back et al. Usable ubiquitous computing in next generation conference rooms: design, architecture and evaluation. In *Ubicomp Workshop 2006*, Newport Beach, CA, USA, Sep 17 2006.
8. Patrick Chiu, Qiong Liu, John S. Boreczky, Jonathan Foote, Don Kimber, Surapong Lertsithichai, and Chunyuan Liao. Manipulating and annotating slides in a multi-display environment. In *Human-Computer Interaction INTERACT '03: IFIP TC13 International Conference on Human-Computer Interaction*, 2003.
9. Brad Johanson, Greg Hutchins, Terry Winograd, and Maureen Stone. Pointright: Experience with flexible input redirection in interactive workspaces. In *Proc. ACM Conference on User Interface and Software Technology (UIST2002)*, pages 227–234, Paris, France, 2002.
10. Meredith R. Morris, Kathy Ryall, Chia Shen, Clifton Forlines, and Frederic Vernier. Beyond "social protocols": multi-user coordination policies for co-located groupware. In *CSCW '04: Proceedings of the 2004 ACM conference on Computer supported cooperative work*, pages 262–265, New York, NY, USA, 2004. ACM Press.
11. Thomas Heider, Martin Giersich, and Thomas Kirste. Resource optimization in multi-display environments with distributed grasp. In *Proceedings of the First International Conference on Ambient Intelligence Developments (AmI.d'06)*, pages 60 – 76, Sophia Antipolis, France, September 19 - 22 2006. Springer.
12. Thomas Heider and Thomas Kirste. Automatic vs. manual multi-display configuration: A study of user performance in a semi-cooperative task setting. In *Proc. HCI'07: HCI.but not as we know it*, Lancaster, UK, September 3–7 2007.
13. Fred D. Davis. Perceived usefulness, perceived ease of use, and user acceptance of information technology. volume 13:3, pages 319–340. MIS Quarterly, Sep 1989.