Joe Tullio
Statement of Research

OVERVIEW

I am a researcher in human-computer interaction interested in augmenting mainstream applications with intelligent systems, in particular those that support communication and coordination between coworkers. Research in context-aware computing and computer-supported cooperative work, along with advances in statistical modeling techniques, have led to a new and active area of research in systems that provide the ability to forecast information about users and share it with others. This information may include a user's presence, availability, future activity, or workload. My research explores the design, implementation, and evaluation of these systems by combining user-centered design principles with a strong technical background to develop prototype systems to support communication and coordination practices. Using qualitative and quantitative methods, I then study how people support and evolve their practices with these systems. My objective is to let users benefit from intelligent support while maintaining a sense of trust and control in their applications.

PREVIOUS WORK: AUGMENTING GROUPWARE WITH FORECASTING

Past work in my graduate career concerned the development of two experimental shared calendar systems named Ambush and Augur. I first designed the Ambush system [5] to perform forecasting on the attendance patterns of a single user, then to share that information with colleagues. Ambush augmented the typical group calendar system with a Bayesian network that predicted attendance at future events. The system learned attendance habits over time and visualized its predictions as bar graphs indicating the probability of attendance. Calendar information was obtained transparently by harvesting an existing PalmOS calendar during synchronization. The Ambush system showed the potential for predicted calendar information to support individual and group work, and to provide contextual data to ubiquitous computing applications. This work also led to several visualization concepts examining how to best convey these predictions in more or less detail.

Following the Ambush prototype, I directed the design of a predictive calendar system that would be widely deployable to a larger group of users and capable of capturing use patterns over time. I added new capabilities such as simultaneous viewing of multiple calendars and identification of common scheduled events among a group of users. The resulting system, named Augur [2], included several improvements over its predecessor. Text-matching algorithms identified those events that had been separately scheduled by different users, and a support-vector machine provided the ability to classify calendar events into various types (classes, seminars, meetings, etc.) and locations. This design allowed colleagues to see which scheduled events they shared with one another and how likely they were to attend them. Augur obtained calendar information from a larger range of existing scheduling tools, including Outlook, iCal, and Mozilla. I oversaw a redesign of Augur's user interface that included views of potential attendees at events as well as their likelihoods of attending. Users could select colleagues' schedules and view them side-by-side with the primary user's calendar. Architecturally, I restructured the system as a secure web service using a combination of JSP, DHTML, and Javascript to generate a user interface compatible with several popular web browsers.

CURRENT WORK: EVALUATING FORECASTING TOOLS IN THE FIELD

One goal for Augur was to be a deployable system capable of facilitating a long-term field study involving dozens of users. Recently, my work has involved the widespread deployment of Augur within an academic setting over an entire semester. I recruited 28 participants to publish their calendars online, and an additional 30 non-publishing participants, for a total of roughly 60 people with accounts on the system. In addition to logging event accesses, logins, and predictions, I conducted a series of four hour-long interviews with a diverse subset of the participants that included students, faculty, and staff. These interviews explored how participants used the system, how they regarded it with respect to other scheduling tools, their degrees of trust in the predictive algorithms, and their comfort with sharing predicted information among different hypothetical classes of viewers. Results indicate that the predictive capabilities of Augur impacted calendar use. Inferences about which colleagues had scheduled the same events had the unexpected effect of encouraging exploration and organizational learning, while predictions of attendance showed sporadic utility in facilitating communication between users. In addition, some participants used the system to diagnose and manipulate the predictions as a means of ensuring privacy and managing their image to other participants. These results imply that designers need to be sensitive to social behaviors such as impression management in future systems.

The Augur field study also raises issues of how shared, machine-generated predictions affect privacy. Key to the success of systems like Augur is an awareness of what vulnerabilities may arise upon deployment, and an ability to maintain this awareness as the system evolves with its users. I have been working with a colleague in the field of software engineering to explore methods for analyzing how presence forecasting can alter the privacy implications of mainstream applications such as group calendar systems. Our technique first uses a structured goal analysis to identify both user and system goal hierarchies, then employs heuristics that examine how information is exchanged and used in the achievement of these goals. In this way, designers can identify high-level privacy vulnerabilities at an earlier stage in the system's lifecycle. Our analysis identified potential vulnerabilities ranging from security of authentication, personal schedule information, and predicted information to the loss of control over presentation through automatic predictions and surrogate representations such as icons. In addition to our own analysis, we conducted a classroom study where HCI students evaluated Augur using this technique. Results show that the most frequently cited issues concern the integrity of attendance predictions and the possibility of automatically matching colleagues with events they have not scheduled. In addition, the study identifies a need for awareness of the predictive algorithms and control over the information they generate.

RESEARCH STYLE

My research methods focus on novel integrations of technologies to produce systems that address observed needs. I combine a strong technical background with an ability to identify research opportunities using ethnographic techniques. My design methods often employ collaborators from other disciplines as well as a thorough grounding in relevant literature. Once implemented, I evaluate my prototypes using both quantitative and qualitative data collection to arrive at conclusions about the system's effects on work practices, communication patterns, privacy concerns, and social dynamics.

Research based in actual practices and behaviors

The difficulties of deploying software to support shared work are well-documented, and the addition of intelligent support has the potential to make this process even more daunting. By combining observational study with a grounding in ethnographic literature for a particular domain, I am able to build systems that support existing practices while adding experimental components that test potential benefits and social effects.

Novel integrations and system-building

A significant portion of my Ph.D. career, as well as my prior work as a software developer, has been spent designing and building systems for deployment. I employ the necessary tools, libraries, and data structures from a multitude of sources in order to execute the design. In addition, I employ novel combinations of technology when it seems appropriate for the task. For example, the Augur system integrated a SQL database, third-party Bayesian network and support-vector systems, the Java servlets API, the Apache Tomcat servlet container, and JavaScript and DHTML-enabled content. In addition to leading the development of Augur, I have developed software for Georgia Tech's Augmented Office project [4], an internal mobile computing venture at Intel's Microprocessor Research Labs, shared large displays at Accenture Technology Labs [3], and the Attentional UI project at Microsoft Research.

Multidisciplinary research: literature and collaborators

From the beginnings of my research career as a staff member of the University of Virginia's Department of Neurological Surgery, I have sought the counsel of collaborators from multiple disciplines and delved into their relevant literature in order to gain insights on my work, to avoid retreading existing paths of research, and to explore methods and theories previously unknown to me. Early collaborators included researchers in perceptual psychology and computer graphics, as well as surgeons and residents. Today, the direction of my research requires grounding in such diverse fields as organizational behavior, software engineering, intelligent systems, and engineering psychology, as well as human-computer interaction. I believe that as computing technology continues to permeate more facets of modern society, any research investigating the effects of this trend will increasingly converge across multiple disciplines, necessitating interdisciplinary collaboration to fully understand the discoveries made.

Qualitative evaluation to get the whole picture

Quantitative data obtained from logs and numeric scores is invaluable for measuring the amount of use a system sees, the accuracy of its predictions, and the subjective ratings of users with respect to various criteria (comfort levels, ease of use, etc.). However, my research stresses the importance of qualitative data in research and relies on this data to gain a clearer picture of why a system is used or neglected and in what situations it becomes more or less useful. Through interviews and observation, I am able to study effects on practice and social concerns, attitudes toward the system, and the degree to which users understand the system's reasoning. By collecting this data over time, I can study these factors at different stages during deployment, and by using a diverse range of users, I can see how effects compare between subjects of differing occupations or between differing inter-subject relationships.

FUTURE DIRECTIONS

For CHI 2004, I co-organized a workshop to examine the nature and extent of research in presence and availability forecasting systems [1]. This field has seen a substantial amount of research activity in terms of architecture, algorithms, and application design since the Ambush system was presented in 2001. Major research issues include treating status such as availability as a negotiated process, understanding sharing preferences for predicted information, and providing clear explanations of models. Evaluation is still a critical need for many existing systems, and I see a number of interesting directions in the near-term for evaluating these kinds of tools. Further, results from my current work indicate additional design considerations that may need to be included in future forecasting systems.

Depending on the algorithm used, users can have markedly different mental models of how an intelligent software component makes its inferences. People are especially prone to error with respect to probabilistic and Bayesian concepts. An important step towards the acceptance of useful intelligent components into everyday applications is the study of how these internal understandings and biases may apply in user interfaces. By eliciting and evaluating these mental models, I can see how intelligent systems are understood and work to provide better explanations and visualizations, and controls to users. Conversely, by experimenting with different intelligent algorithms, I can determine how the interfaces to such algorithms shape mental models and relate to the tasks performed. I have recently begun work in eliciting mental models of Augur participants, which results showing that higher error rates and less control over input may have the effect of producing overly elaborate and incorrect mental models.

Along these lines, it is important to consider the image presented by group applications of their users to one another. This becomes even more important when machine-generated information, such as inferences about activity or presence, is shared among a workgroup. Many research opportunities exist in terms of evaluating how users manage their image through shared applications, and what facilities can be provided to facilitate this management, especially in the case of forecasting systems.

In the longer term, I see the potential for intelligent systems to become pervasive in everyday group applications. Without understanding the unique social as well as technical challenges presented by such integrations, however, designers risk undermining their users' senses of control and trust, ultimately impeding a system's chances for adoption. Through my research, I hope to develop techniques for building and maintaining trust in intelligent systems so that people may reap the benefits they can offer.

REFERENCES

1. Tullio, J., Begole, J., Horvitz, E., Mynatt, E. Forecasting presence and availability, *ACM Conference on Human Factors in Computing Systems Extended Abstracts (CHI 2004)*, pp. 1713-1714.

2. Tullio, J., Goecks, J., Mynatt, E., Nguyen, D. (2002) Augmenting Shared Personal Calendars. *ACM Symposium on User Interface Software and Technology (UIST 2002)*, pp. 11-20.

3. Huang, E., Tullio, J., Costa, T., McCarthy, J. (2002) Promoting Awareness of Work Activities Through Peripheral Displays. *ACM Conference on Human Factors in Computing Systems Extended Abstracts (CHI 2002)*, pp. 648-649.

4. MacIntyre, B., Mynatt, E., Voida, S., Hansen, K. M., Tullio, J., Corso, G. (2001) Support For Multitasking and Background Awareness Using Interactive Peripheral Displays. *ACM Symposium on User Interface Software and Technology (UIST 2001)*, pp. 41-50.

5. Mynatt, E. and Tullio, J. (2001) Inferring Calendar Event Attendance. *ACM Conference on Intelligent User Interfaces (IUI 2001)*, pp. 121-128.