

Applying Commonsense to Augment User Interaction in an Intelligent Environment

Fahim Kawsar¹, Mostafa Al Masum Shaikh², Tatsuo Nakajima¹, Sheikh I. Ahmed³

¹Department of Computer Science, Waseda University, Tokyo, Japan

²Department of Information and Communication Engineering, University of Tokyo, Tokyo, Japan

³MSCS Department, Marquette University

{fahim,tatsuo}@dcl.info.waseda.ac.jp, mostafa_masum@ieee.org, sheikh.ahamed@mu.edu

Keywords: Commonsense, User Interaction, Augmented Artefacts.

Abstract

In this paper, we present a system that augments user interaction to elicit intelligence of a living room adorned with aware artefacts. These artefacts are computationally augmented everyday objects, like a phone, a light, a TV, etc., which are enabled to communicate their operational states among themselves. The operational state of an artefact is usually changed to another state as a result of an interaction by the user. Our system augments such interactions by improvising other spatially correlated artefacts with the help of commonsense knowledge. For example, if a user picks up a ringing phone while the TV is on, the system either mutes or reduces the TV volume. In this case the phone's state is changed to "in use" state due to deliberate interaction of the user and simultaneously the state of the TV may be changed to "mute" or "low volume" by the system. Such idea of augmenting user interaction has been elaborated in this paper. We present the computational model and the implementation detail of our approach and also demonstrate its feasibility through an informal user study.

1 Introduction

We envision that with the proliferation of low-cost sensors, smart artefacts and spontaneous communication technologies pervasive applications will find a universal place at our home and will transform the physical space into an intelligent one. Consider the following scenario:

"Chantelle recently moved into a new home. After moving she visited her neighbors to make their acquaintance. They had an incredible home. When she sits on the couch of the living room, the air-conditioner and the lights automatically went on. While they were talking the volume of the TV faded away. After they stopped talking and started watching a movie, the blinds came down and the lights dimmed when the movie started and they went on when it ended."

The intelligent home like Chantelle's neighbors usually model situations to improve user experience. Our home

artefacts have many built-in functions and many of these functions and corresponding operational states are related to each other. Understanding the correlation among the operational states of multiple artefacts poses an interesting opportunity to improve user experience. Often we usually interact with multiple artefacts together or in sequence to accomplish an activity, like muting the TV after picking a phone call or turning on the light after opening the door etc. If the artefacts states can be connected to a knowledge-based network, part of these sequences can be automated by understanding the user's primary interaction and by improvising other artefacts states utilizing commonsense knowledge. In this paper we present living room artefacts that are computationally augmented and can share their operational states with each other [5,7]. A commonsense knowledge base and user preferences are applied to improvise artefacts states. Automating interaction sequence to improve user experience is an interesting research topic that has drawn significant attention in recent times. Specifically, the context-aware computing observes human activities and situational contexts to understand users intention to provide just-in-time services [1,2,4,13]. However, instead of recognizing human activities entirely, our approach solely considers the operational states of artefacts and the correlation among them. To understand the correlation of artefacts states, we have developed a commonsense knowledge base from an empirical study on living room interactions. The knowledge is represented as a graph where the nodes indicate the states of the artefacts and the edge between two nodes indicates the correlation of the two states happening together or in sequence. When an artefact state is changed due to deliberate interaction of a user, this graph is used to improvise related artefacts states. This automation is further overlaid by users preferences. Informal user study of our prototype system revealed satisfactory user acceptance where 78% of the time the system was able to correctly augment users primary interactions.

In the subsequent section, we present the computational model that is the basis of our system. We then proceed to the design issues of our approach followed by the description of the prototype implementation. We then present the feasibility of our approach and position our

research with respect to the related work before concluding the paper.

2 Computational Model

Our basic goal is to detect a user's primary interaction with an artefact and to change the states of other spatially co-located artefacts automatically as a natural reaction of the user's primary action. To formulate our approach, we define a physical space in the following manner:

A space W is a container of a collection of networked and aware artefacts A_i , and is represented as i.e., $W = \sum_{i=1}^n A_i$

where each artefact A_i possesses a set of operational states $S_{A_i} = \sum_{j=1}^{m_{A_i}} s_{A_i,j}$ and a set of properties

$P_{A_i} = \sum_{k=1}^{q_{A_i}} p_{A_i,k}$. The space W is associated with a

hierarchical location model and the member artefacts of the corresponding space can reside in any of the location within this model. Considering at a certain time a user is capable to interact with a part of the physical space and with a few artefacts, we define a space (i.e., W) in terms of several active zones (see figure 1). This division can be fine grained by applying egocentric situative model [12].

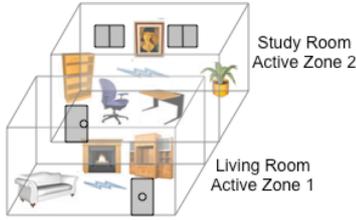


Figure 1: Physical Spaces with Multiple Active Zones

For the sake of simplicity here we define an active zone E_t at a certain time t as a subset of W where the member artefacts are in a particular state. It is represented as:

$$E_t = \forall_i \forall_A (\exists_A location(A,l) \wedge \exists_{s_A} state(A,s))$$

Where the predicate $location(A,l)$ indicates that an artefact A is at location l and the predicate $state(A,s)$ indicates that an artefact A is at state s . When a user interacts with an artefact in E_t , that artefact's state is altered causing the active zone to move into a new state E_t' . Our goal is to improvise this E_t' further by identifying the spatially co-located artefacts and their states that can be switched to other appropriate states as a consequence of the user's interaction in E_t . We represent the augmented active zone as E_t'' . Therefore, to transit

from E_t' to E_t'' (as shown in figure 2) we consider the semantic relationships among the states of the co-located artefacts obtained from commonsense knowledge base and user's preference profile. In the next section, we discuss the design rationales that convert this computation model into a working prototype.

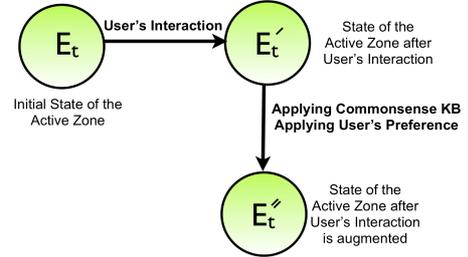


Figure 2: State Transitions of Active Zones

3 Design Issues

The computational model presented in the earlier section poses us three design challenges.

Augmented Artefacts and Location Management: Everyday artefacts like a phone, a desk, a lamp, etc., should be capable of understanding and sharing their state-of-use. To enable this, everyday artefacts were augmented with sensors and actuators. Each of the artefacts is self-contained, connected to a network and capable of sharing $state(A,s)$ predicate. These artefacts are developed following the design guidelines proposed by Kawsar et al. [5]. The locations of the artefacts were maintained centrally following a lightweight location system *Spreha*, where the static artefacts are used as reference points for tracking the mobile artefacts [6]. This location system essentially provides the $location(A,l)$ predicate.

Commonsense Knowledge Base: To find the semantic relationships among the states of the artefacts, i.e., which of the operational states might occur together or in sequence a knowledge base is required. Typically in an intelligent system, intelligence comes from the designer which often fails to match end users expectations [3]. An alternative approach is to observe the target users' common practices empirically and to model those practices in the system. In addition, human commonsense can improve machine intelligence to enrich users' experiences [10]. We have decided to adopt the later approach and generated a commonsense knowledge repository based on an empirical study on living room interactions of 5 households over one week. In the next section we will elaborate our data collection phase and the knowledge representation method.

User Preference: Every user has his/her own understanding and perspective towards an application and wants to personalize it regardless of its proactive behavior [3]. Hence for the success of an application, it is essential to allow end users to personalize the behavior of intelligent applications. Here by personalization, we mean the active

participation of end users to customize the adaptive behavior of the system. In our approach, after the system deployment end users are asked to personalize each of the proactive actions utilizing a speech interface. Users' preference values are then overlaid with the system defined intelligence obtained from commonsense knowledge base. In our current prototype multiple users are not considered, thus the preference conflict issues are out of scope of this paper. In the next section, we will explain this preference handling mechanism in detail.

Following the aforementioned design guidelines, we have developed a prototype system that is explained in the next section.

4 System Description

In the current prototype system, we have augmented several artefacts of a living room to model a typical living room scenario. Figure 3 shows the component architecture of the system.

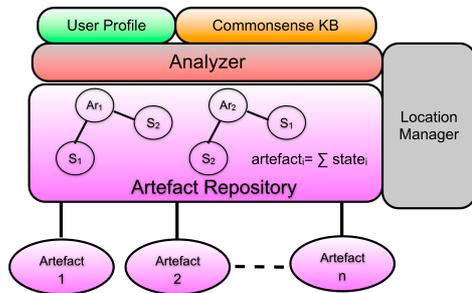


Figure 3: Overall System Architecture

4.1 Artefact Repository

This component represents the underlying environment. Each of the artefacts are computationally augmented to be aware of their current state of use and connected to a network. Whenever an artefact's state is changed, the corresponding artefact notifies the Artefact Repository. In our current prototype scenario six artefacts are deployed. Their computational augmentation and state details are shown in table 1.

Table 1: Augmented Artefacts and their States

Artefacts	Augmentation	States
Door	2D Accelerometer	Open, Close
Light	X10 Module	On, Off, Dim
TV	Simulated TV application running on a PC	Off, Mute, On-Volume
Couch	Pressure Sensor	In use, Just Used
Phone	Pressure Sensor	Picked, Just Used
Coffee Maker	Light Sensor	In Use, Just Used

We have followed the design principle proposed by Kawsar et al. [5] to represent these augmented artefacts.

Cookie sensor node [9] and Gumstix¹ are used for instrumentation. Sensor data are mapped to operational states locally at artefact end applying simple threshold analysis and are broadcasted to the Artefact Repository accordingly when the state is changed. Figure 4 shows the artefacts and their augmentation platforms that are used in the implemented prototype.

4.2 Location Manager

In our system we have used a centralized location manager utilizing the lightweight artefact based location system *Spreha* that uses static artefacts as a reference point to track the location of mobile artefact applying triangulation on Bluetooth signal strength [6]. Door and Couch are used as static location reference points to track the other four artefacts. Since Cookie sensors are packaged with Bluetooth radio, *Spreha* provides an economic and feasible solution for tracing spatially co-located artefacts.

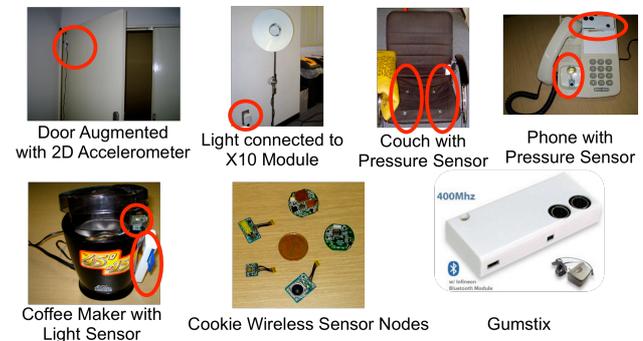


Figure 4: Augmented Artefacts and their Platforms

4.3 Commonsense Knowledge Base

As mentioned in the design section, we have opted to build a knowledge repository to generate the semantic relationships among the states of the multiple artefacts. In order to build the knowledge base we have taken an empirical approach by video taping the living room interaction of five households over a week. Three of the households have two or more members. Participants were selected randomly through an open invitation in a social networking website and were promised to keep their identities anonymous. It was also ensured that all the households' living rooms contain the six artefacts that are finally developed. For participating in the study they were compensated with a gift. Analysis of the video clips exposed some typical interaction pattern in a living room: For example, from interaction records the following patterns are detected:

After stepping into the room he turns on the light.

After sitting on couch he turns on the TV.

While watching movie he turns on the air-conditioner.

Before leaving the room he switches off the lights.

Similarly, another clip analysis revealed the following:

After sitting on sofa she makes a phone call

After stepping into the room opens the blind

¹ <http://gumstix.com/>

Before watching TV she closes the blind

Interestingly, we have observed a similar pattern of artefacts usage in all the households and several interaction cycles occurred repeatedly in multiple households. If we look at the interaction and usage sequences, it essentially shows the pattern that we all are familiar with and which in other words can be called commonsense [8]. By analyzing such pattern of interaction of users with living room artefacts, we have classified the relationship of artefacts states into two groups: “Should Have”, and “Good To Have”. A conditional probability is utilized for this classification and is calculated as follows:

If an artefact A_i reaches a state $s_{A_i,j}$ and another artefact A_k reaches a state $s_{A_k,j}$ immediately or in sequence and if A_i and A_k are spatially collocated, i.e.,

$$location(A_i,l) = location(A_k,l)$$

then the conditional probability of having the two states together or in sequence is,

$$P(s_{A_k,j} | s_{A_i,j}) = \frac{P(s_{A_k,j} \cap s_{A_i,j})}{P(s_{A_i,j})}$$

where $P(s_{A_i,j})$ is normalized to the total number of occurrences of state $s_{A_i,j}$ of artefact A_i .

Should Have relationship between two artefact states signifies that occurrence of this sequence of interaction has a probability value higher or equal to 0.5

Good To Have relationship between two artefact states signifies that occurrence of this sequence of interaction has a probability value less than 0.5.

Hence, by empirical analysis of the video footage and calculating the conditional probability values we have developed the common sense knowledge base or augmenting users interaction in a living room. The commonsense knowledge of interaction is represented as a connected graph where each node indicates a state of a particular artefact and the edge type indicates either of the two relationships and the edge value is the conditional probability of the two states happening together or in sequence. Figure 5 shows the graph that represents the commonsense knowledgebase for the six augmented artefacts that we have used in our system.

This knowledge base is used by our system to augment user interaction. i.e., when user interacts with an artefact, the system improvises the spatially co-located artefacts state by applying the rules. For example the value (conditional probability) of the edge connecting node *Door <Open>* and *Light <On-Bright>* is 0.9 and the edge is of type *Should Have*. Therefore whenever the door is opened, the system automatically turns on the light.

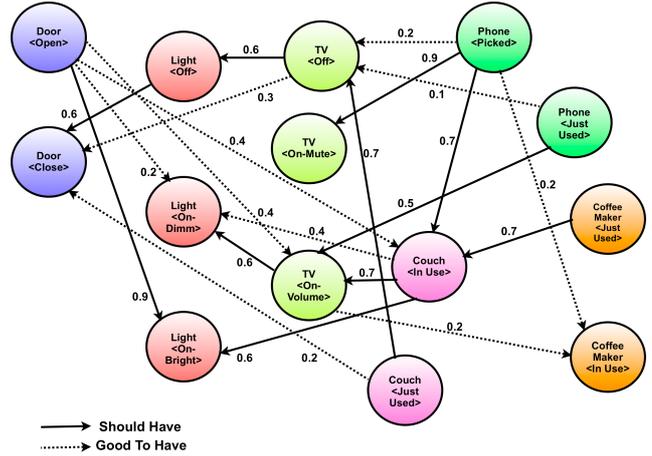


Figure 5: Connected Graph representing the Commonsense Knowledgebase. Edge value represents the conditional probability of two states happening together or in sequence

4.4 User Profile

During the system deployment, the knowledge base depicted in figure 5 serves as the default search space to find the possibilities of appropriate augmentation in the active zone. However, before employing any auto augmentation, end users are asked for providing their preferences regarding the proactive action. The User Profile component takes care at this stage. To interact with the system, the end users are offered an audio module composed of a speech recognizer and a text-to-speech component (see figure 6). Users’ responses are only considered in binary “Yes/No” manner for simplicity and tagged with each action. To enable the audio module a speaker and microphone are deployed in the living room test-bed. A preference file is maintained which contains the user preferences and the basic rules obtained from the knowledge graph is overlaid with these preferences before invoking the automatic action of the system. The overlay algorithm is mentioned in the next subsection.

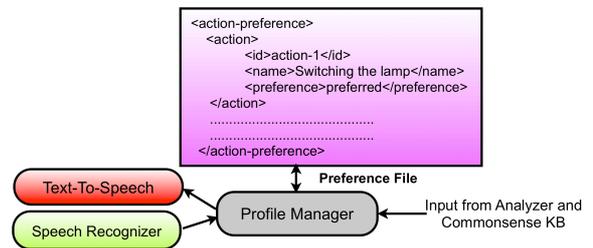


Figure 6: Internal Architecture of User Profile

4.5 Analyzer

This component agglutinates all the system components. Artefact Repository manages and holds all artefacts states. When an artefact state is changed, the Artefact Repository notifies the Analyzer. Analyzer then consults the Location Manager to identify the active zone of the primary artefact and the co-located artefacts within the same active zone.

The Commonsense Knowledge Base is then consulted if there are multiple artefacts in the active zones to associate with the primary artefact's state. If an association is found from the commonsense knowledgebase then the relationship is considered. If the relationship is "Should Have" then the corresponding actions are invoked by the system. However, if the relationship is "Good To Have" then the user profile is consulted to see the user preference. If no user preference is found regarding the action then user is asked using the text-to-speech component of User Profile Manager. Users input is captured through the speech recognizer and the preference file is updated. If a user prefers a "Good To Have" action then the knowledge is updated for that user to reflect users preference and the corresponding relationship is changed to "Should Have". Considering there are primarily two artefacts (TV and Light) in our test environment that can be actuated, examples of the analyzer actions are:

Turn on the light when a user steps into the room
Mute the TV when a user picks the phone
Turn off the TV when user leaves the room
Switch the light to dim mode when user turns on the TV
Turn off the light when the user leaves the room

In the next section we report the informal end user evaluation of our system performance.

5 Evaluation

To evaluate the system's performance and acceptability, a user study with 15 peoples is conducted.

5.1 Experiment Procedure

Initially we introduce our research agenda to the participants. Then we invite them to our living room test bed to interact with the artefacts. Each trial took about 30~40 minutes. After that, participants were presented a questionnaire followed by an interview. The participants' demography is mentioned in table 2.

Table 2: Demography of the Participants

Participants	Age Range	Profession
10 male	23~35	Graduate Student, Lawyer, IT Professional, Researcher.
5 Female	25~52	Graduate Student, Business Consultant, Researcher.

5.2 Guiding Topic

We have designed the questionnaires and the interview following the guideline topics mentioned below:

- **Functional Features:** Whether the participants like the proactive functions of our system.
- **Interaction and Personalization:** Whether the participants are satisfied with the interaction and personalization features.
- **Overall Unobtrusiveness, Usefulness and Usability**

5.3 Quantitative Results

The first issue is the accuracy of the interaction augmentation. In the user trials there were 381 primary interactions with the six artefacts by the 15 participants where there were rules for augmentation. Out of these 381 incidents, 298 times (78%) the interaction was correctly augmented, meaning the system could change the state of an artefact autonomously considering another artefacts state. We consider this result promising. Next, we present the end user' assessment towards the system. Since some of the questions were very direct in the form of affirmative and negative response, we have requested the participants to rate their answers on a 5-point scaling. In the following figure 7 results are shown which highlights that the proactive actions by augmenting user interaction and the personalization feature of our system were appreciated. The rationales behind their likings are explained in the next subsection. There were mixed reactions on voice-based interaction, specifically some users told us that it is completely unrealistic to be verbose in a living room.

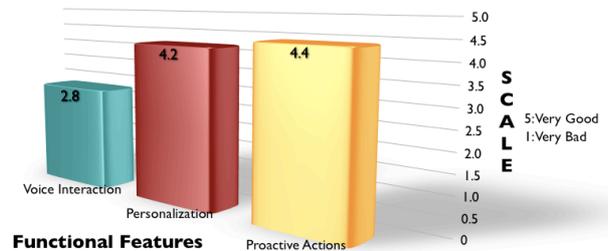


Figure 7: Users' Responses on Functional Features

5.4 Qualitative Results

Figure 8 shows the overall user response on the quality aspects of our system. We consider the responses are quite promising. Interview with the end-users revealed several issues: users considered our system useful because it allows interacting with multiple artefacts that are physically apart. Also, the semantic relation of the artefacts and user interaction seems natural to the end users. For example, muting or reducing volume during a conversation or turning on the light after stepping into room is a natural interaction that user are familiar with and are already available commercially in numerous offices. Similar actions in household are quite acceptable to them. Also, our proactive actions were quite limited which users accorded warmly since they consider too much automation is distracting. Furthermore, since our system provides control to the system to personalize the actions, users were willing to accept our solution since they considered it is important for them to be in the control of the system. The graphs showed in figure 7-8 effectively reflect these views. A negative aspect of the system is the speech-based interaction that they found disturbing and annoying most of the time and suggested alternative mechanisms.

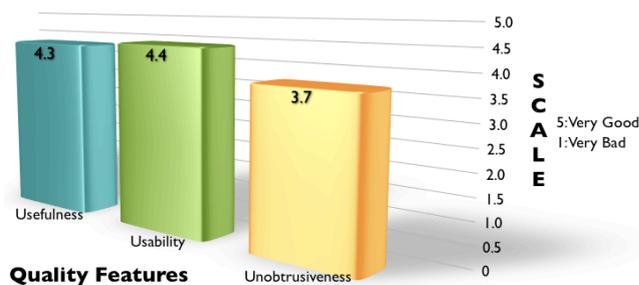


Figure 8: Users' Responses on Quality Features

6 Related Work

Considering the work presented in this paper primarily targets augmenting user interaction to form an intelligent environment, we look at the alternative approaches that have been explored for implicit interaction. There is a rich body of literature in the context aware computing domain where users' operational contexts are observed to provide just-in-time services. Primary approach that has been investigated is the human activity recognition by machine learning algorithms either utilizing large number of scattered sensors deployed in the physical space or using wearable sensors. Sensor data is then extrapolated with other situational contexts like time, identity, location, users' preferences etc. to model user centric contexts. For example, Tapia et al. used Bayesian Classifier to recognize activity of users by modeling sensory output collected from scattered sensors deployed in the house [4]. Schmidt uses the term *Implicit HCI* to extend users primary interaction by understanding her perception towards environment primarily utilizing users context collected from sensors [1]. Other researchers in pervasive computing domain have tried similar approaches. Recently Patel et al. have shown that how residential power line events can be classified using machine learning algorithms to predict users' activities [11]. From commonsense-based system perspective, Lee et al. presented augmented kitchen application where sensory data is overlaid with commonsense to assist user in kitchen activities [2]. In all these or similar works, recognition of the user activity using sensors is primarily used to provide contextual services. Our approach differs from them considering we only look at the semantic relationship of the artefacts states. Although we have augmented our artefacts with sensors, we have not attempted to recognize human activities explicitly. Instead a commonsense knowledge base is used to form a semantic relationship of artefacts states. This relationship is used to augment a user's interactions with the environment in a contextual manner.

7 Conclusion

In this paper, we present an integrated approach of combining aware artefacts with commonsense knowledge to provide a proactive intelligent space. Our system can augment a user's interaction with the environment by analyzing the correlation among the states of augmented artefacts. The primary contribution of this work is the

utilization of semantically related artefacts states to form an intelligent environment. Instead of recognizing human activity by scattered sensors, we solely focus on artefacts states. The secondary contribution of our approach is the combination of commonsense and user preference to generate a rule base intelligent system meshing physical artefacts, which effectively highlights the application of integrated intelligence. We also reported an informal user study of our system that raised several interesting issues for further exploration. One important avenue of our future work is to integrate more aware artefacts in the environment to scale our system in more complex scenarios. We are also considering the personalization aspects in a multiuser environment. We hope to report some exciting results on these issues soon.

References

- [1] A. Schmidt, "Implicit Human Computer Interaction through Context." *Personal Technologies* 4(2), 2000
- [2] C. J. Lee, L. Bonanni, H. Lieberman, H. Espinosa, and T. Selker, "Augmenting Kitchen Appliances with a Shared Context using Knowledge about Daily Events." *IUI 2006*.
- [3] D. Norman, *The Design of Future Things*. Basic Books. 2007
- [4] E. M. Tapia, S. S. Intille, and K. Larson, "Activity Recognition in the House Using Simple and Ubiquitous Sensors." *Second International Conference on Pervasive Computing 2004*
- [5] F. Kawsar, K. Fujinami, and T. Nakajima, "Experiences with Building Intelligent Environment through Sentient Artefacts." *The 3rd IET International Conference on Intelligent Environment (IE07)*. Ulm, Germany, 2007
- [6] F. Kawsar and T Nakajima, "A Lightweight Indoor Location Model for Sentient Artefacts using Sentient Artefacts." *The 22nd Annual ACM Symposium on Applied Computing (SIGAPP)*, 2007
- [7] F. Kawsar, K. Fujinami and T. Nakajima, "Augmenting Everyday Life with Sentient Artefacts." *The 2005 Joint Conference on Smart Objects and Ambient Intelligence, sOc-EUSAI '05*. 2005
- [8] H. Lieberman, A. Faaborg, H. Espinosa, and T. Stoky, "Commonsense on the go." *BT Technology Journal* 22, no. 4. 2004
- [9] K. Hanaoka, A Takagi, and T Nakajima. "A Software Infrastructure for Wearable Sensor Networks." *The 12th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications*. 2006
- [10] P. Singh, B. Barry, and H. Liu, "Teaching Machines about Everyday Life." *BT Technology Journal*. 2006
- [11] S.N. Patel, T. Robertson, M. S. Reynolds, and G. D. Abowd, "At the Flick of a Switch: Detecting and Classifying Unique Electrical Events on the Residential Power Line." *9th International Conference on Ubiquitous Computing*, 2007
- [12] T. Pederson and D. Suiire., "Towards an Activity-Aware Wearable Computing Platform Based on an Egocentric Interaction Model." *Fourth International Symposium on Ubiquitous Computing Systems*. 2007