DETECTING PRIVACY IN ATTENTION AWARE SYSTEM

Jérôme Maisonnasse, Nicolas Gourier, Oliver Brdicka, Patrick Reignier and James L. Crowley

INRIA Rhône-Alpes, France

Abstract: Privacy and security are important issues for acceptance of ambient technology. For this reason, we propose a detector of relations between users and objects based on attention aware system. This detector is inspired by social sciences and completes the representation–driven model mainly used in ontology-based approaches. We discuss the role of context in these approaches and managing privacy. An attention model is derived from the gravitation model and cognitive psychology approaches. This model exploits contextual elements such as position, speed and saliency of objects in a scene to estimate shared attention. An application has been developed to demonstrate the efficiency of this system when managing windows applications, while guaranteeing the security and privacy of users.

Keywords: Privacy, Attention Aware system, Focus of Attention, Context, Ambient system, Intelligent Environment, Video Tracker.

INTRODUCTION

Privacy and security have always been an important issue for acceptance of new technologies. In ambient intelligence, systems should guarantee security for private information in different social situations and assure protection against potential attack of bad intended people. In this work, we focus on the management of devices within an environment. This work is a proposition to integrate more intelligence into intelligent environment. Our approach uses perceptual components and provides information concerning context and activity evolution by detecting relations between humans and objects. We exploit context evolution to delimit public and private areas of occupant of intelligent environment. This is relevant to privacy and can also be introduced in systems with a higher level of control and intelligence.

Recent progress in computer multimodal perception promises new developments in the field of ambient applications and pervasive systems. Such systems aim at offering services by taking into account the current user’s activity in a specific situation. In intelligent environments, more and more devices are capable of perceiving user activity and proposing appropriate services. However, intelligent environment must take care of information overload produced by an unorganized environment, where all devices sue for users anywhere without coordination between them. This kind of behaviour is uncomfortable for users and compromises their privacy. When a system interrupts one user, it is not often this user but all users that are interrupted. In this case, surrounding people become the involuntary witness of hidden information of the user. Thus, addressing the right user at the right moment is essential. This loss of control is not acceptable, especially when private information like administrative information (credit card number, bank number, etc.) become available for anyone who is present at the wrong moment. That’s why we must detect potential users and their connection while doing an activity. Therefore, it is important to detail relations between different users and devices which share a limited space. An intelligent system which addresses a user in the best way by guaranteeing privacy must understand these relationships.

Nowadays, the relationship between users or objects are conceptually defined by ontology or object modelling (4)(3) or implicitly defined or learned by a probabilistic data distribution (1)(10). In this paper, we propose an attention aware system which takes into account relationship between users and contextual elements within the same shared physical space, in order to qualify and to quantify relationships. Qualifying relationships corresponds to specifying who pays attention to whom. Quantifying attention means to compute a value which represents the strength of links between users and objects. Attention is a perceptual interface between the physical world and our representation of world. Attention selects which objects are presents in our mind or not. However, attention is a cognitive filtering process whose capacity is limited and oriented in direction to a subset of physical objects within an environment. The most common metaphor to describe attention is a magnifying glass. The more an object is in the focus of attention, the more this object is deeply analyzed by high level processes that involve memory or others intellectual skills. This process increases or decreases the interest of objects in the high level processing chain. Then, it is essential to define what is context for each user, by considering his attention state. There exist different definitions of context in literature. In
this work, we attempt to build a bridge between a social science definition that we adapt for an ambient intelligent system. We propose a new interpretation of personal context in order to provide detection tools and to identify what elements are common between different users and what elements are specific to one user, so that a computer system can delimit privacy. However, delimiting who is concerned by an activity and where are situated context boundaries is a difficult task. Indeed, when a user is in immediate physical contact with another user, they contribute to the same global definition of the situation (7). By adjusting their personal representations and by integrating contextual events, people create a cognitive shared space where they negotiate a mutual comprehension about current context and current activity. Psychology studies show that human activity is more unexpected than we perceive. Limits between different activities are fuzzy when users share the same physical space (8). Outside laboratory conditions, activity evolves depending on unexpected external events.

We propose an attention model to dynamically detect focus of attention and delimit the subspace of representation which is personal or shared with other persons. An interaction between people occurs when they share the same information (12). A cognitive explanation of how people share information is their ability to produce a mutual intelligibility of the current situation. Perception of current situation is defined by enabled shared resources in a physical, social and cultural environment, more or less stabilized (12). Based on focus of attention detection, an attention aware system could estimate what elements are mutually shared by users and what elements are not shared. However, determining user focus of attention is a difficult task. Focus of attention is an internal cognitive operation which cannot be perceived directly. This perception must be estimated from external observations like head orientation or gaze direction or by recognizing explicit attention events from contextual objects, like devices which determine when a user interacts with them. Computer-human interaction quality will be enhanced, if the system can choose the best place and moment to send a message by taking into account the position of the attention focus of the current and surrounding users in relation with their activities.

In this paper, we model available contextual element in an intelligent environment in order to compute mutual attention of people and objects. Based on this model, we build an attention service providing information about the configuration of attention of surrounding users. A program client installed on a device is connected to the service, and manages different configurations of windows applications when privacy is not ensured.

RELATED WORK

Based on the idea that the social world is organized and understandable in the way action is produced, computer science attempts to extract invariant features to describe activity. Such features are extracted either from perceptual data, e.g. in computer vision (1), or from user interaction with software and devices (9). However human activity is situation dependant and is not only determined by a cognitive plan (15). Some work attempt to integrate contextual elements to understand user activity (3). A context aware system integrates a representation about the environment, like state of devices, positions and movements of occupants. Two concepts of context coexist. The first representational approach is more a computer designer approach. Context is a description of states of the environment and its occupants. This approach is well situated to connect very heterogeneous computing system. Activity takes place inside a context but it does not depend on it. Then context is “any information that can be used to characterize the situation of an entity” (5). Almost all studies based on this approach develop ontology (3) or object model (11) technologies to infer user activity from contextual information. An entity can be a person, place or object considered relevant to user and application. The second approach comes from sociology and is called interactionist approach. Here, the relational property between entities is dynamically redefined by emergent activity. Then, only a subset of properties of context becomes relevant in relation to the activity. Activity is dependant on context and vice versa. Dourish (6) describes four differences with the representational approach: “First, rather than considering context to be information, it instead argues that contextuality is a relational property that holds between objects or activities. It is not simply the case that something is or is not context; rather, it may or may not be contextually relevant to some particular activity.

Second, rather than considering that context can be delineated and defined in advance, the alternative view argues that the scope of contextual features is defined dynamically.

Third, rather than considering that context is stable, it instead argues that context is particular to each occasion of activity or action. Context is an occasioned property, relevant to particular settings, particular instances of action, and particular parties to that action.

Fourth, rather than taking context and content to be two separable entities, it instead argues that context arises from the activity. Context isn’t just “there,” but is actively produced, maintained and enacted in the course of the activity at hand.” (6).

Because, this approach is original in computer sciences and well adapted to integrate unexpected events in the real world, we choose to develop some concepts to enhance comprehension of the human activity by a computer system. In order to integrate the interactionist approach between activity and context into a computer system, we choose to look for a more mechanic theory. A theory that is easily calculable and manageable by a computer system. Psychology offers relevant models to understand how people could interact in relation to contextual elements.

ROLE OF CONTEXT IN THE PRODUCTION OF MUTUAL INTELLIGIBILITY

To explain the way agents are able to communicate, it is necessary to admit that they share mutual knowledge. Theory of mutual knowledge has the characteristic to produce a regression at infinity. But this theory cannot be integrated into a cognitive explanation of production and comprehension of communicative acts. Sperber and Wilson developed a weaker but empirically more adequate concept, the mutual manifestness. For Sperber and Wilson, ‘a fact is manifest to an individual at a given time, if and only if, this individual is able at this time to represent this fact mentally and to accept his representation as being true or probably true’” (13). In other words, a fact is manifest when it has the characteristic to be perceptible or deduced by an agent at a given time. A fact can thus be manifest without being known. However, some facts can be more manifest than others. To model this, we associate a degree of salience to each fact. The salience is a function of perceptual and cognitive capacities of the individual, and of his physical environment. For example, let us suppose that a telephone is ringing in a room where an individual A is sitting at an open window, and that at the same time a car is passing in the street. In this case, it will be strongly manifest for A that the telephone rang, but less clear that a car passed. This is the case because of the difference of salience between the ringing telephone and the car noise, the fact "telephone is ringing" is more manifest, i.e. has
more chance to be perceived or deduced than the fact "a car passed". Sperber and Wilson (13) define the personal cognitive environment (PCE) as whole facts which are manifest for a given individual. A shared cognitive environment indicates all the facts which are manifest to several individuals. From the example of telephone by imagining that another individual B is in the same part as A. In this case, by supposing that they have the same perceptual capacities, it is manifest for A and B which telephone is ringing. This means simply that they are able to perceive or deduce the same fact, and not that they share a belief, a knowledge, or a representation concerning this fact. The Mutual Cognitive Environment (MCE) indicates a shared cognitive environment in which the identity of individuals who have access to this environment is manifest. A and B share a cognitive environment which includes all facts and especially their co-presence. As they share the same environment, they can establish an interaction in relation to their common perception of contextual events. To be even more precise, Salambier (2004) propose to reduce shared contextual (SC) objects thought activity filtering, as in Figure 1. As perception is not a passive system, he proposes that "shared context is a set of contextual information or events mutually manifest for a set actors, at time t, taking into account their perception and cognitive abilities of their tasks which should be made by them". This definition is more precise than Dey's definition of context as "any information that can be used to characterize the situation of entities" (5). Dey does not precise how user information is selected. We define context as the whole set of objects which are manifest for an individual and capable to modify his interpretation of the situation. Context matching with personal cognitive environment describes above. Reasoning about different cognitive environment offers different relevant configuration of context aware systems to address specifically users.

![Diagram of shared context](Image)

**Figure 1.** Graphical representation of relationship between MCE and SC of two users (letter A and B represent two users) (12).

An attention aware system is a particular type of context aware system which reasons on the users’ cognitive environment built from the attention configuration of users. When the personal cognitive environment is delimited, the attention system is able to use different strategies to interact with users in the best way. Personal cognitive environment represent all devices, persons or object which are perceived and mentally represented by one user. This cognitive area is his relevant context for him. Efficiently addressing user with efficiency imply to use objects which are inside personal environment. Addressing two users at the same time implies to pass via objects which are inside the mutual cognitive environment. To be more task dependant, the system can use the shared context. The two previous cognitive environment are typically identified because privacy is not guaranteed. However, privacy can not be defined for only one user. An ambient intelligent system should be able to manage privacy of a users group in relation to others groups. In the field of privacy management, maybe, the more relevant cognitive environment, added by ourselves, is private cognitive environment (table 1). This environment contains the whole set of relevant objects perceived by only one particular user. The cognitive environment should directly concern directly almost applications of privacy management. The system is able to address this user discreetly by guaranteeing that anyone else could access to private information.

**TABLE 1.** Cognitive environment and different strategies to select objects.

<table>
<thead>
<tr>
<th>Cognitive Environment</th>
<th>Object Selection Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal Cognitive Environment</td>
<td>Subset of object manifest for one person</td>
</tr>
<tr>
<td>Mutual Cognitive Environment</td>
<td>Subset of object manifest for two persons</td>
</tr>
<tr>
<td>Shared Context</td>
<td>Subset of object manifest for many person and directly relevant for current task.</td>
</tr>
<tr>
<td>Private Cognitive Environment</td>
<td>Subset of object manifest for one person and not manifest for other</td>
</tr>
</tbody>
</table>

**ATTENTION MODEL**

The cognitive mechanism evoked above by Sperber and Wilson is called attention process. Context is thus defined throughout an attention filtering process and is the result of a combination between user activity control process and environment stimulation control process. These two antagonist control processes enable someone to stay concentrated in a task in noisy environment and to react correctly if a relevant or sudden event appears. The nature of these two processes recalls notions presented in the interactionist approach. Context is thus dynamically defined by association of current activity and environment setting modification. That is why, attention model is an interactionist approach. Defining attention is very difficult task. For this reason, we develop a model motivated by the gravitational model. This model includes some interesting concepts which recover attention properties. First, we present a short recall about what is the Gravitation Model. The first Law of Universal Gravitation has been formulated by Isaac Newton in the 17th Century. Any two objects in the Universe exert gravitational force on each other, with the universal form (1). This force is proportional to the product of their masses and inversely proportional to the square of the separation between the two objects. An example is shown in Figure 2.
In this work, we compute focus of attention to delimit context boundaries for each user and to detect when people share same resources or not, on the basis of their position and the salience of contextual elements.

The focus of a person is defined by the direction of attention which is the combination of its external and internal factors. The external factor of a person is determined by the attraction coming from other people, objects or artefacts which inhabit the environment. We adapt the gravitational model to simulate a person’s attraction towards other persons or objects. Each person or object has a salience \( m \). The salience corresponds to the mass in the gravitational model. We suppose that salience is invariant, which allows computing the attraction vector of each person towards the people and the objects in the environment using the gravitational model. The salience could be defined on perceptive, social or situation features. The internal factor of a person is determined by the person’s current goal or current activity, regardless of its environment. This factor could be assimilated to intentionality concept. Cues of intentionality of a person are for example current speed and gaze direction or interaction with an object. The internal factor can also be represented by a vector. For the moment, we just take into account the current speed. An important intention vector could be perceived an important directed concentration to an object used a task. Both external and internal factors vectors are combined so that the influence of the external factor decreases exponentially with the internal factor, as shown in Figure 3. For our application, the influence of the attraction becomes negligible when the speed is higher than the top speed \( v_{\text{max}} \) of human running of 4 m/s. In this case, we suppose that only objects present in the direction of the run are relevant for the runner. We compute the attention vector as a linear combination of internal and external factors.

\[
\vec{F}_{\text{attr}} = \sum_{j \neq i} \vec{F}_{i \rightarrow j} = \sum_{j \neq i} \frac{G \, m_i \, m_j}{r_{ij}^2} \, \vec{u}_{i \rightarrow j}
\]

The Fundamental Principle of Dynamics (3) enounces that the derivative of the quantity of movement of an object \( i \) is equal to the sum of the forces exerted on this object. By supposing that the mass \( m_i \) of the object is constant, we can compute the acceleration of this object (4).

The acceleration \( \ddot{a}_i \) of object \( i \) stands for the attraction of other objects on object \( i \). In particular, it reflects the fact that objects with little masses are more attracted by objects with bigger masses than objects with bigger masses towards objects with little masses.

\[
\frac{d^2 \vec{r}_i}{dt^2} = \vec{F}_{\text{attr}}
\]

\[
\ddot{a}_i = \sum_{j \neq i} \frac{G \, m_j}{r_{ij}^2} \, \vec{u}_{j \rightarrow i}
\]

In this work, we compute focus of attention to delimit context boundaries for each user and to detect when people share same resources or not, on the basis of their position and the salience of contextual elements.

The focus of a person is defined by the direction of attention which is the combination of its external and internal factors. The external factor of a person is determined by the attraction coming from other people, objects or artefacts which inhabit the environment. We adapt the gravitational model to simulate a person’s attraction towards other persons or objects. Each person or object has a salience \( m \). The salience corresponds to the mass in the gravitational model. We suppose that salience is invariant, which allows computing the attraction vector of each person towards the people and the objects in the environment using the gravitational model. The salience could be defined on perceptive, social or situation features. The internal factor of a person is determined by the person’s current goal or current activity, regardless of its environment. This factor could be assimilated to intentionality concept. Cues of intentionality of a person are for example current speed and gaze direction or interaction with an object. The internal factor can also be represented by a vector. For the moment, we just take into account the current speed. An important intention vector could be perceived an important directed concentration to an object used a task. Both external and internal factors vectors are combined so that the influence of the external factor decreases exponentially with the internal factor, as shown in Figure 3. For our application, the influence of the attraction becomes negligible when the speed is higher than the top speed \( v_{\text{max}} \) of human running of 4 m/s. In this case, we suppose that only objects present in the direction of the run are relevant for the runner. We compute the attention vector as a linear combination of internal and external factors.

\[
\vec{F}_{\text{attr}} = \sum_{j \neq i} \vec{F}_{i \rightarrow j} = \sum_{j \neq i} \frac{G \, m_i \, m_j}{r_{ij}^2} \, \vec{u}_{i \rightarrow j}
\]

The Fundamental Principle of Dynamics (3) enounces that the derivative of the quantity of movement of an object \( i \) is equal to the sum of the forces exerted on this object. By supposing that the mass \( m_i \) of the object is constant, we can compute the acceleration of this object (4).

The acceleration \( \ddot{a}_i \) of object \( i \) stands for the attraction of other objects on object \( i \). In particular, it reflects the fact that objects with little masses are more attracted by objects with bigger masses than objects with bigger masses towards objects with little masses.

\[
\frac{d^2 \vec{r}_i}{dt^2} = \vec{F}_{\text{attr}}
\]

\[
\ddot{a}_i = \sum_{j \neq i} \frac{G \, m_j}{r_{ij}^2} \, \vec{u}_{j \rightarrow i}
\]
The first component is a speech detector which indicates when someone starts or finishes speaking. The detector is based on an artificial neural network technology (16). The second component is composed of a video tracker which detects any entities in the environment and converts their positions from picture scale to environment scale (2). An example can be seen in figure 6.

The third component is a detector of keyboard or mouse events. When a user interacts with the personal computer where the component is installed, a XML message is produced. These three perception components send their information into the fourth component. Attention services computes attention state of each user from its input components. It returns an XML message which describes how attention resources have been distributed for each object. A fifth component is installed on a personal computer and process only attention information concerning it. Then, by parsing this message, this component becomes aware of owner attention defined by an id and surrounding people. Therefore, the attention controller application can act on window environment. The user put over a label “private” on a window application and if a security rule is violated then window is minimized. This rule is an acceptation test whether an object is inside the private cognitive environment. This means that only one user pays attention to this object. A threshold on attention resources attributed to a particular object delimits if the object is accepted or not into a cognitive environment. In the same application, one icon indicates the state of attention of surrounding users in relation to the computer. Based on three different thresholds, three colours are used to inform users about interest of surrounding people: red colour indicates that someone really ventures to access to your private information, yellow colour indicates a lower probability that someone accesses to your private information and green colour indicates that the situation is secured (figure 7).

Some example situations are proposed to detail more deeply how attention model integrates input information and derives attention state of each user. In these situations, two persons work in their office. The first person enters in the room (A) (situation 0) and is walking in direction of this personal computer (0) (situation 1). He starts to interact with his personal computer to realize different tasks (situation 2). After some minutes, the second user joins him by entering in the office (situation 3). After greeting the first user (situation 4) who has interrupted this work in order to answer, the second user walks in direction to his personal computer and the two persons start to working together (situation 5). Then, the second user decides to stop to work and approach close of the first user without inform him (situation 6). To understand how the attention model works, we present results corresponding to these six different situations in the next table 2.

<table>
<thead>
<tr>
<th>Results of Attention Model</th>
<th>Comments about model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Situation 0</strong></td>
<td>When the first person (2) enters in office, the environment is only composed of three object. We can add more device like telephone if this last could be used. The personal computer (1) and (2) are switch on but anyone interact with its, then their saliency is low (weight = 0.1). User attention resource is only defined by his initial direction speed provided by video tracker. This direction is interpreted like an intention. For this reason person (2) pays little attention to computer (0) represented by black colour pie.</td>
</tr>
</tbody>
</table>

...
Situation 1
When person (2) walks, its current speed increases. Then intentionality increases too with regard to speed. In the direction of attention vector (black line) two personal computers are present. Each of them is represented by black and blue grey pie in attention available resources. The angle formed by attention vector and the vector which links person (2) and computer (0) is more narrow than the other angle of the personal computer (1). Thus more attention resources are attributed to computer (0) than computer (1).

Situation 2
When person (2) starts to interact with his personal computer, its salience increases from 0.1 to 0.5. Attention vector is thus essentially computed on attraction of the computer because speed is close to zero. In this situation, all attention resources are directed to computer (0).

Situation 3
Nothing changes for the person (2). But for the newcomer (3), his attention is little catch by the presence of the first user and his personal computer. In this situation, the quantity of attention to personal computer is too low to trigger the attention controller reaction. Computer (0) stays in the private cognitive environment of the first user (2).

Situation 4
The second person greets the first person. Then his salience increases to 10, because speech is an attracted modality. The first person (2) stops his work, is clearly attracted by speech. Therefore, personal computer (0) salience decreases because person (2) does not interact with it. For user (3) quantity of attention is high enough to trigger attention controller which minimizes private application windows with the “private” tag.

Situation 5
Each user works on his computer. The salience of computers increases to 0.5 with mouse and keyboard interaction. Each user is mainly concentrated on his personal computer. This fact is represented by a large attribution of attention to computer (0) and (1).

Situation 6
The second person (3) approach discreetly near person (2). This proximity is detected by attention service which warns person (2) by red colour signal and by minimizing private windows application.

For this application, the salience of personal computers and users are defined exponentially from prototypical situation as it is represented on the figure 8.

CONCLUSION
We presented an approach completing the representational context. Based on a knowledge database, the representational approach is able to infer some relevant situations where users are secured or where privacy area is compromised. Our method is perception dependant and completely dynamic. This approach is more situation dependant and near of interactionist approach. An attention aware system does not work with semantic information but with numeric values. This aspect is specific to our approach in comparison to semantic approaches. Attention services could be used like a simple detector for other intelligent components. In our application, this aspect is demonstrated by the attention controller. Therefore, cognitive environments seem to be a relevant concept to control an attention aware system. The different cognitive environments could be required to divide context in different subspaces and thus offer different levers to manage services in relation to current activity and surrounding people.

We have developed a windows manager application called attention controller. This application works in real time. However, it is dependant on video tracking performance. The, quality of our model can be increased by using a more robust system of localization. Very few input data are used. These input data combined perception like position of people and device information. The combination of these two kinds of input increases precision of the attention process. However, we can imagine to integrate others perceptual components like head pose estimation in order to increase the quality of the internal factor. In fact, head pose is highly correlated (14) with attention direction and could be a useful cue. However, attention and resources distribution can not be limited by head pose. The reason is that people are able to share their attention in parallel, like speaking with someone during an interaction with a computer. In this case, the attention is shared between these two entities. We need to correct some limitations of the attention model which could return a wrong attention distribution. Basing only distribution of attention resource on angle formed by attention vector and vector which links a person with on object product sometimes errors when more objects are present in the scene (figure 9). We foresee to correct these errors in a further works.
Figure 9. User (1) pays more attention to users (4, dark green) and (5, light green) than the closest users (0, black) and (2, green).

The first results are very encouraging and we hope to develop a complete attention aware system which manages devices not only for privacy issue but to coordinate interaction between users and digital environment too. A systematic evaluation of the attention model is in progress in order to validate the model-driven selection process in comparison to the performance of a human supervisor.

REFERENCES


