Towards a Trust Analysis Framework for Pervasive Computing Scenarios

Michael Butler, Michael Leuschel,
Stéphane Lo Presti
University of Southampton
SO17 1BJ
Southampton, United Kingdom
{mjb, mal, splp}@ecs.soton.ac.uk

David Allsopp, Patrick Beautement,
Chris Booth, Mark Cusack, Mike Kirton
QinetiQ ltd
WR14 3PS
Malvern, United Kingdom
{d.allsopp, Patrick.Beautement, cjmb,
cusack,
kirton}@signal.qinetiq.com

ABSTRACT

We present a scheme for highlighting the trust issues of merit within pervasive computing, based on an analysis of scenarios from the healthcare domain. The first scenario helps us define an analysis grid, where the human and technical aspects of trust are considered. The analysis is applied to a second scenario to examine its suitability. We then discuss the various categories of the analysis grid in the light of this examination and of the literature on the subject of trust. We believe that this approach could form the basis of a generalised trust analysis framework to support the design, procurement and use of pervasive computing.

Keywords

Trust analysis framework, pervasive computing, healthcare scenario, trusted agent

1. INTRODUCTION

The increasing availability of small footprint computing platforms (PDA, mobile phones) and the deployment of next-generation wireless networks (UMTS , Wifi, Bluetooth) are helping to make the vision of pervasive computing a reality. Pervasive computing [1] places the user at the centre of an environment populated by services accessible through devices embedded in physical objects. In contrast to the current mode of human-machine interaction, where all tasking is performed by the user, pervasive computing seeks to soften the attention and learning requirements for a user within the environment, by enabling the system to make inferences about the user's needs.

Many observers believe that agent-based computing is an important enabling technology for pervasive computing [17, 12]. It is envisioned that collaborating agents, capable of describing, discovering and accessing services, will assume important information and service management roles in future pervasive systems. To enable agents to reason about the capabilities of the services on offer, it is anticipated that they may exploit the common understanding of terms, relations and services across communities promoted by the Semantic Web or Grids. At the human-machine level, interface agents may interact with humans to elicit and report information, and to provide an atmosphere of ambient intelligence [8].

The application of pervasive computing to healthcare is attracting a great deal of interest at the moment [10, 2]. The potential monitoring, reasoning and reactive capabilities of agent-based pervasive environments could reduce the cognitive and physical burden on healthcare professionals, and improve the quality of care and standard of living of vulnerable patients [3]. A key point in the acceptance of this paradigm is its applicability, which can be demonstrated by realistic scenarios. Pervasive computing scenarios have already been designed [8]. It is critically important that these scenarios are validated by subject matter experts, so that they plausibly depict people and processes within the healthcare domain. A key principle for pervasive computing design should be to fit the technology to the task, rather than the opposite.

A central aspect of pervasive computing is the notion of trust [2]. Since pervasive computing focuses on the user, technical features such as security (which is often confused with trust) are no longer sufficient to correctly design and implement distributed systems. The subjective concept of trust not only enables users to better understand pervasive computing, but also opens new ways of solving existing problems, such as security [15], management of online communities [16, 7] or automated negotiation. The literature on the subject of trust, though extensive and with a wide range of applications from sociology [5] to computing [13], teaches us that the foundations of trust are not completely understood and its ramifications are deep in many systems. Though trust classifications already exist [4], there is no clear consensus on the definition of this concept, partly due to the fact that this definition depends on the context of use.

We believe that agent technologies, which embrace the subjective and uncertain aspects of trust, will be of particular importance in pervasive environments. In some instances, agents within the system will effectively become

6^{th} *International Workshop on Trust, Privacy, Deception, and Fraud in Agent Societies* (AAMAS'03)

extensions of the person, and must be given a human-like capability to reason about trust. We are looking at techniques that involve agents observing, modelling and learning from the behaviour of other parts of the system over time. This information is then used to establish trust relationships in an analogous way to humans, in order to procure and offer pervasive healthcare information services seamlessly. For determining how we can apply agent technologies that promote trust in pervasive computing, we need to fully understand the trust issues of significance within potential pervasive healthcare applications.

In this paper we describe a methodology for highlighting the key trust issues within pervasive healthcare scenarios. Matters of trust in distributed computing are often discussed in terms of abstract concepts or security features, and it is sometimes difficult to appreciate the impact of particular trust issues on the users of the system. Because of the human-centric nature of pervasive computing, it is critically important that trust is explored from the user's perspective. The analysis scheme reflects this imperative by considering the trust issues from the standpoint of the user.

To aid an understanding of trust matters, we have developed a number of plausible scenarios which contain use-cases that highlight the interactions between a user and her pervasive environment. These scenarios form the foundations of our methodology and we believe that their development and analysis can provide a valuable holistic view of trust in pervasive computing.

In section 2, we describe our trust analysis scheme, that is detailed in the subsequent sections. Sections 3 and 5 present summarised versions of two pervasive healthcare scenarios, each of which has been validated by healthcare experts. Section 4 summarizes the results obtained from the examination of the trust issues of the first scenario in the form of an analysis grid. This analysis grid is then used on the second scenario in section 6. Section 7 relates the various categories of the analysis grid, tying them to concepts found in the literature, and exhibits the key points and flaws of the analysis. We conclude by summarising our results and presenting future work.

2. TRUST ANALYSIS SCHEME

Starting with a pervasive computing scenario, our trust analysis scheme involves iteration over four steps. The first step involves the validation of the activities, events and participants within the scenario by a subject matter expert. For the first scenario, presented in section 3, clinical procedures were validated by a speech and language therapist. It is critically important that the scenarios accurately reflect the way in which people would use the pervasive technologies to support them in their work.

The second step in the scheme involves the extraction and examination of use-cases to determine the trust issues. These use-cases are extracted from vignettes within the scenario, each of which may highlight one or more trust issues. For extracting trust issues from the use-cases, the reviewer identifies what she believes to be a trust issue, and illustrates it with one or more concrete examples taken from the scenario. The analysis in terms of use-cases helps to bound the search for trust issues to a certain context.

In the third step, the initial examination of trust issues undergoes peer review, cross-checking and classification. Peer review supports the extraction of additional trust issues from the perspective of another potential user, who may have a different vien on trust issues. The trust issues are cross-checked against other scenarios, to establish whether common trust themes exist in different application areas. These common themes form the basis of our categorisation of the trust issues. We assume that the generalisations derived from the trust analysis are plausible because they have been derived from the user's interaction with the system.

In the fourth step, the scenario is refined by adding new use-cases, or removing existing ones. We view a scenario as a living document, whose purpose is to provide a framework in which to illustrate possible applications of pervasive computing, and to extract the most relevant trust issues. It is important that the scenario reflects the trust concerns of all the stakeholders involved, and it should be updated to represent different priorities. Finally, the updated scenario is validated by the experts who first validated it, thus confirming the assumed plausibility of the generalisations.

Next, we present the first scenario which describes the visit of a speech and language therapist.

3. FIRST SCENARIO: THE SPEECH AND LANGUAGE THERAPIST

Janet is a speech and language therapist who is visiting Ella, a three year-old, who was referred by her health visitor. The links between primary healthcare information systems are well established, and the health visitor's referral was sent to the local Speech Therapy Service automatically, encrypted and digitally signed. On arrival, the Speech Therapy Service's triage agent made the decision to assign Ella to Janet's caseload, based on conversations with Janet's personal agent regarding her current workload, and the initial assessment of the health visitor. The referral contains information represented in Semantic Web ontologies that the triage agent uses to reason with. The health visitor has recognised that Ella is slow in her language development, and is mostly using vowel sounds only. Janet has a good record in treating young children with this condition, and the triage agent has learnt this.

In order to make the appointment, the triage agent sends Ella's case notes and the referral to Janet's personal briefing agent, who oversees her diary. The triage agent and Janet's briefing agent belong to the same agent security domain, which has a security policy of allowing unfettered access to any other services in the domain, after authentication of identity certificates.

Every morning Janet's personal agent briefs her on the day's cases. The form of the briefing depends on a number of factors, such as time, location, available interface devices, and also Janet's preferences, which were determined by user modelling and machine learning techniques. Janet prefers to be briefed about cases during her car journey to work, or on the way to visit clients. The personal agent conducts the briefing by placing a call to the hands-free 3G mobile phone in her car. Speech recognition over a limited vocabulary allows Janet to identify herself to her agent, and to control the structure and the tempo of the briefing. The personal agent uses a combination of GPS, together with information from mapping and traffic monitoring Web Services, to issue verbal directions to Ella's house.

Prior to Janet's visit, her personal agent broadcast a discovery request across the Speech Therapy Grid, requesting information on possible courses of action for delayed speech and language development. The Grid forms a virtual department spanning speech therapy departments across the country. All departments are members of the same security domain and can exchange information in the context of the domain's trust policy. The treatment strategies returned as a result of the search request were processed and then prioritised based on success rate.

On arriving at Ella's house, Janet's personal agent downloads the case notes to her wireless-enabled PDA, via the car's 3G mobile link. This highly sensitive information is provided to the PDA in an encrypted form on a limited time lease, for the estimated duration of the consultation unless the lease is renewed by a user-sanctioned action. The case notes can only be accessed via the exchange of identity certificates, stored and delivered by Janet's smartcard.

During the consultation, Janet uses her PDA to record speech assessment results. This information is stored and later forwarded to the Speech Therapy Service to be attached to Ella's records. Ella has feeding difficulties that were not previously identified, and Janet is not trained to deal with this particular condition. The parents seem particularly anxious about Ella's development, and it is important that they receive reassurance and advice during the course of the home visit. Janet decides to seek a second opinion.

Janet asks her personal agent to obtain additional information on which to make an informed decision. Janet qualifies the request she made prior to the visit, given Ella's additional feeding difficulties. The agent joins the appropriate security domain and generates a treatment strategy based on the case studies available. The search results are de-personalized, and then sent encrypted over a fast connection to Janet's car, and then on to her PDA.

After the consultation, Janet leaves Ella's house and drives away to the next home visit. The end of the session is inferred from the change in location. Information such as the length of time for the consultation is recorded and will be used to support waiting list statistics and for generating efficiency metrics. Additionally, Janet does not have the burden of performing administrative tasks, such as filling in timesheets and performing mileage calculations, since these are compiled continuously throughout her working week.

4. TRUST ANALYSIS GRID

The second step in our scheme begins with identifying vignettes and use-cases where trust issues arise in the scenario. Example vignettes comprise Janet's briefing by her personal agent and the use of the Speech Therapy Grid. Use-cases are then identified among the vignettes and examined to highlight the trust issues. Further examination reveals that the trust issues can be grouped into categories, each one based on a concept. These concepts are abstract, yet anchored in concrete examples from the scenario, and shape some facets of the concept of trust. By increasing the level of abstraction, we define a first trust analysis grid that will serve to analyse other scenarios. This trust analysis grid is a dynamic tool that will evolve while we progress along our analysis scheme.

In the following, the trust categories are presented in the form of an analysis grid. Each category is identified by the name of its concept, and is then defined and illustrated by one example from the scenario. The grid is presented in a compact form here.

Source vs. Interpretation

An interpretation of the data is less trusted than the source data itself. For example, sound and video recordings of the consultation may be considered a more trustworthy information than the interpretation that Janet gives in Ella's records.

Reliability

This property indicates that a service operates according to its specification. In the scenario, the network reliability enables Janet to trust her Speech Therapy Service and the Speech Therapy Grid.

Reasoning

Each participant manipulates the data to process it, in order to make decisions or answer a request. This process can weaken the trust a participant has in the system. For example, Janet would distrust the triage agent if it directly modifies her agenda, bypassing Janet's briefing agent.

Personal Responsibility

A person must remain responsible for the actions she performs, since they are not mediated by a trusted system. The property of accountability is important to put a significant level of trust in the system. For example, the system cannot prevent Janet from reviewing case notes in the car, if she chooses to do so regardless of the illegality.

Authorization

Any agent accessing a piece of information must have the permission to do so. In the scenario, Janet can change some of the information from Ella's case notes but not all of them.

Identification

Identity is important to differentiate the participants and communicate with one of them. Janet is probably identified before accessing her PDA, for example by biometrics (finger print), or by a smartcard.

Privacy

This property ensures that the personal information of a user is not accessed (and hence used) without him knowing it. The results from the Speech Therapy Grid are de-personalized to ensure this property.

Integrity

This means that the data is free from unauthorized manipulations. The referral's digital signature is a means to prevent a malicious agent from modifying it without detection.

Usability

This aspect of trust encompasses various elements, like the intrusiveness of the mechanisms used to interact with the user, or its usefulness. The way Janet's personal agent briefs her shows some of these elements.

Audit trails

An audit trail lists all the actions performed, together with their parameters (e.g. identity and authority of the performer) and these information should not be modifiable. Ella's referral is an example of such audit trail.

Accuracy

The more information that is detailed, the more precisely trust can be evaluated in the system. This is illustrated when Janet seeks the help of the Speech Therapy Grid to give a more accurate answer to Ella's parents.

Harm

This aspect goes hand in hand with trust, since trust is a belief, and it may be misleading and harm the system. For example, if the trust that Janet has in her the triage agent is low, she must be prepared to loose some time to readjust the wrong assignments it may give her.

After the second step of our methodology, we present in the next section a different scenario that will serve as a basis for applying the analysis grid.

5. SECOND SCENARIO: A PRE-HOSPITAL INCIDENT

As John leaves his house driving to work, his cell phone is aware of the car, and switches to automatic answer mode. The phone is able to take the usual text and voice messages, and also to give specific information to certain callers. Meanwhile, the car makes use of the national traffic monitoring system to see how fast the roads are flowing. The car can thus detect when traffic conditions cause an unexpected delay, enabling John's diary agent to attempt to reschedule any important meetings likely to be missed.

The traffic is bad today, due to road works. The morning winter sun is blinding, and the car darkens the top of the windscreen, as John prefers. People seem to be driving worse than usual and too close to one another. The head-up display gently indicates this, but John can't find space to move into.

John puts on the stereo using the car's speech recognition. Brake lights flare and John is jolted alert. There seems to be a wall of slowing cars and smoke is pouring from the wheels of the car in front. Before his foot even hits the pedal, the collision avoidance system applies the brakes. John's car was too close to avoid a collision, but at least he slowed before glancing off the lorry on his left and hitting the car in front. As the motorway grinds to a halt, it appears that three cars have crashed. One of the cars skidded trying to avoid a tyre on the carriageway, and was struck by the car in front of John. Other motorists have managed to avoid the initial accident, but some have had minor collisions.

The emergency services already know much of the situation. As soon as the cars' airbags were triggered by the crash, the cars transmitted a distress call, including their location and the number of occupants (detected by simple pressure sensors in the seats). The crash barrier has been bent out of shape by several vehicles, and a fibre-optic sensor line confirms to the emergency control room that an accident has happened. Other motorists will begin to phone 999 soon and provide confirmation, and extra, though probably confused, information. The national traffic monitoring system will be notified of the blockage, and will begin to re-route traffic to other roads, lessening the congestion around the accident. The first car's phone was too badly damaged to transmit its call, but for 999 calls it was able to piggy-back on the phone of the second car using short-range networking.

The emergency control room dispatches a small number of police, fire and ambulance vehicles immediately. The incoming calls from other motorists, and images from a traffic camera on a nearby bridge, seem to confirm the seriousness of the accident, and further vehicles are dispatched. The information known so far is shared between all of the vehicles en route. Information on traffic flow and speed is also shared between the vehicles to enable them to avoid blocked or slow routes. The dispatch and arrival of the vehicles is logged automatically to provide statistics on response times.

The controller is concerned about the availability of Accident and Emergency (A&E) beds. The nearest hospital seems to be at capacity, and the next nearest can only take three casualties. The hospitals will be asked for firmer estimates as soon as actual casualty numbers are known.

The traffic police are first on scene, and begin making the area safe with cones, and securing the crashed vehicles as best as they can. The video feed from their speed camera is available to the control room, but it doesn't provide much useful information for the bandwidth consumed. A still image is shared with the vehicles en-route though. The police confirm the number of vehicles involved, and the number of casualties. They quickly take a few evidential photos of the scene, and begin basic first responder treatment. Policemen are unsure how to handle one of the casualties, who appears not to be breathing, as they are worried about worsening any spinal injury present. They are connected to the nearest ambulance, and are talked through the appropriate treatment by the paramedic.

One of the policemen is trying to hand John over to the first paramedic on scene, who is from the fire service, but the policeman is told to keep holding John's head still while the paramedic triages the other casualties. He informs the ambulances (and the emergency control room) of his findings by radio. The emergency control room enters this information into the log for the incident, which is shared with the receiving hospitals.

The ambulances are arriving on scene and, after checking with the fire-fighters that the scene is secure, the paramedics continue treatment. They record their assessment and treatment onto normal paper report forms, but these are backed by smart clipboards that record and recognise the handwriting and ticked boxes. Each patient is given an RFID tag, normally on a wristband, to enable the incident record to follow them around the system.

The patient with breathing difficulties is causing concern and he is being transported to hospital immediately. A paramedic is assisting breathing with a bag-valve-mask and oxygen, and so she is not able to fill in the report forms as she goes along. However, the data from diagnostic equipment, such as the pulse oximeter, can be logged and sent ahead. The paramedic has control over this if she wishes, to compensate for false readings or inaccurate data. She fills in the report form when she gets the chance, and this will be available immediately via the patient's RFID tag.

John seems to be relatively unhurt, but he is immobilised with a cervical collar and board until spinal injury can be ruled out. The spinal board, sadly, doesn't have any sensors yet, but it does have an embedded RFID tag to identify which ambulance trust it belongs to. The fire-fighters are busy cutting up the car in front, and one of them is taking a few quick photos with his helmet camera for the incident support crew, to give them some idea of the scene.

The nearest hospital confirms that it is now able to take two casualties. The next nearest hospital can take another two, which will be sufficient. The hospitals are able to view the photos and other information if they wish, to assess mechanism of injury, and to negotiate who should get which patients, if particular expertise or facilities are required.

6. ANALYSIS OF THE SECOND SCENARIO

We now use the trust analysis grid defined in section 4 to examine the second scenario. This implies systematically searching for vignettes and use-cases corresponding to the various categories of trust issues. This third step of the analysis scheme is intended to give us insight in the various categories of the analysis grid, their relative importance and their relationships. We only give here some of the use-cases identified by applying the analysis grid, while respecting their relative proportions and aggregating some of them for the sake of simplicity.

Source vs. Interpretation

- Paramedics' on-scene assessments (hand-written and captured by the clipboards) are interpretations.
- The data coming from the diagnostic equipment is a source data, whereas the information controlled by the paramedic is an interpretation.

Reliability

- The phone service should correctly take and reply to text and voice messages.
- The car equipments (collision avoidance system, airbag system, distress call transmission) must be reliable.
- The emergency controller actions (forward information to hospitals, connect the police and the paramedics, notify of the availability of hospitals' A&E beds) must be performed correctly.

Reasoning

- John's diary agent reschedules John's meetings.
- The emergency controller decisions coordinates the various information (car transmissions after airbag activation, fibre-optic sensor line from crash barriers, motorists' phone calls) and consequently dispatches the police, fire and ambulance vehicles.
- The paramedics triage the casualties by evaluating their criticality.

Personal Responsibility

- Drivers are accountable for their actions during the accident.
- Policemen must make basic first responder treatment, but they must contact the paramedics for a casualty they are unsure of prior to any action.
- The paramedic who is assisting breathing is responsible for not altering the integrity of the data from the diagnostic equipment.

Authorization

- The sharing of information between the vehicles (traffic flow, speed) and with the hospitals is authorized by the controller.
- The paramedic assisting the patient with breathing difficulties can control the diagnostic equipment's data after unlocking the equipment with his ID card.
- The doctors assigned to the casualties are authorized to treat their patients, e.g. according to their CVs.

Identification

- Every person involved in the incident can be identified: the wounded via their RFID and policemen, firefighters, paramedics, and doctor via their institutional ID.
- The person calling John must be authenticated by his phone answering machine.

Privacy

- John's phone answering machine must not give information about John to the callers it responds to.
- The controller's log must be protected against information leakage by an authorization system.
- Incident records can only be used among the participants managed by the controller and for a short period of time after the incident.

Integrity

- The police cameras' video feed must not be altered.
- The paramedics' assessments are made via the paper and electronic reports and those two systems must be kept consistent.
- Data from the diagnostic equipment is altered by the control of the paramedic.

Usability

- John's car system takes care of darkening windscreen, measuring the distance to the surrounding cars, recognising his voice commands, and controlling the brakes so that he is not surprised by their emergency action.
- The paramedics' smart clipboards release them from typing notes.
- While assisting the patient with breathing difficulties, the paramedic has her hands full and should not be interrupted by any system (that is why she is not obliged to fill in the report).

Audit trails

- John's estimated time of arrival to his office.
- The actions of the cars during the accident.
- The controller's log, comprising times taken by the vehicles to arrive on scene, number of A&E beds at the hospitals, patients' incident records, controlled data from the diagnostic equipment, fire-fighters' quick photos, and every access to the log with its reasons.

Accuracy

- The amount of cars with minor and major collisions, with information on the number of occupants.
- Resolution of the video feeds and quick photos.
- Number of casualties indicated by the police and hospital's A&E beds.
- Scale of the data from the diagnostic equipments.

Harm

- John's callers not having the right information can lead him to miss deadlines or lose information.
- Information not up-to-date (e.g. number of A&E beds) or inaccurate (e.g. video feed) can lead a participant to evaluate incorrectly a parameter and take a wrong decision (e.g. taking care of a patient already treated, free an allocated A&E bed, etc.).

7. REFINING THE TRUST ANALYSIS GRID

In order to finish the third step of our analysis scheme, we try to relate the various categories of the trust analysis grid. We complete this refinement by comparing the trust analysis grid with related work.

7.1 Relation between the categories

All the categories composing the trust analysis grid appear in the second scenario, with less use-cases for the categories Source vs. Interpretation, Identification and Harm, and most for Usability, Audit trails and Accuracy. Some use-cases appear in several categories at the same time and some of these combinations appear several times. Examining these combinations in the light of the literature on the subject of trust, we relate and simplify the categories from the trust analysis grid.

Authorization, Identification, Privacy, and Harm

Authorization mechanisms may or may not use Identification, as the second scenario shows for RFID tags or the phone piggy-back system. Traditional distributed systems implement security as an authorization mechanism combined with an authentication mechanism using some form of identity. Blaze et al. [11] demonstrate that identities are not necessary by using credentials binding public cryptographic keys to authorization to perform a task, rather than binding keys to names. Rasmusson and Jansson [9] give the motivations behind these situations: anonymity is a means to ensure objective reviewing of an agent's competence whereas identification allows to support recommendations.

Privacy is implicitly tied with the notions of Authorization and Identification, as we can clearly see in the second scenario with John's phone answering machine and the patients' incident records. Privacy can be seen as a by-product of the authorization mechanism or as an empowerment of the user [18]. As clearly underlined in the literature, the loss of privacy can have serious consequences, especially in an healthcare environment [10].

Furthermore, most of the use-cases of the category Privacy are included in those of the category Harm, but the opposite is not true, as the example of the controller's decisions suggests. Loss of privacy leads to harm for the concerned individual, though it may be authorized for example when an authority investigates a case on the individual. Therefore, Privacy can be included in the category Harm, or similarly Harm may be decomposed into more atomic elements, including Privacy.

Privacy and Identification are two topics that have important roles in of human societies, notably though laws and regulations. As distributed systems are evolving to provide better services to users, we can see that these two categories are more closely tied to the systems architecture. Nevertheless, Harm seems to be a more significant issue than Privacy and brings an interesting direction for further research. For example, trustworthiness can depend on the existence of a credible social or economic threat [14]

Reliability and Integrity

Reliability plays the same role for services or equipments as integrity for information or data. They complete each other in that they ensure that every process and data is correct according to its definition. In the example of the second scenario, the controller is reliable only if it ensures the integrity of the data it is managing.

Source vs. Interpretation, Integrity, and Accuracy

These three categories are properties of the data used in the scenarios. Although losing any of them may have the same consequences, these categories are generally not equivalent, since their variations may be independent. For example, if the controller's input parameters are less accurate, it may manage the emergency situation less efficiently, but it will still rely on correct data to make sound decisions.

These three categories can be quantified, though no standard measures of each of them exists. These values may serve as basis of a metrics of trust, such as those found in systems like Advogato [16] or the Semantic Web [6].

Audit Trails and Personal Responsibility

Audit Trails log actions, their authorizations, and the parameters of these elements (temporal constraints are sometimes imposed in order to respect privacy). Audit Trails are tied with Personal Responsibility in that they are used to prove it. If a responsibility is fulfilled according to an audit trail, it credits its subject with some reliability, while if it is not fulfilled, sanctions can be applied to harm its subject. In the scenario, the actions of the paramedics are logged so that, if a casualty deteriorates, the cause of the problem can be traced and, if a paramedic made an error, responsibility can be determined.

7.2 Summary of the refinement

After the third step of our analysis scheme, few simplifications of the trust analysis grid have been found. Only Privacy seems redundant with Harm, and Usability and Reasoning have few links to other categories. The links between the various categories sketch a taxonomy of the trust issues as follows:

- **Subjective categories**
  The categories Reasoning, Usability and Harm share a fundamental property of trust: subjectivity [5]. They involve the agent's internal state and knowledge and express its beliefs. These categories also provide part of the context that is used to interpret trust relationships.
  These categories require further investigations as they relate to well-known subjects in the literature: recommendations and reputations for Source vs. Interpretation, game and risk-management theories for Reasoning and Harm. Few work has been done so far on Usability in the perspective of trust [13].
- **System categories**
  The categories Authorization, Identification, Personal Responsibility, Reliability, and Audit Trails relate to the underlying system in the various scenarios. This system may be a device, a computer program, or the socio-economic system.
- **Data categories**
  The categories Source vs. Interpretation, Integrity, and Accuracy describe the data from the point of view of the trust issues.

7.3 Comparison with related work

Our analysis share some similarities with the requirements stated by the TRUST-EC project [19]. We discovered that the category Availability was missing in our grid. This category brings a new light to some vignettess of the scenarios. For example in the first scenario, the piggy-back system relies on the availability of phone systems in the surrounding cars, which cannot be assumed in general. On the other hand, our trust analysis grid adds the categories Source vs. Interpretation, Reasoning and Usability. This points to the fact that many of the trust methodologies tackle the problems from a technical point of view, rather than a human-centric one. Thus, many of the subjective facets of trust are evaded, but these concepts are more directly applicable to real-world applications.

8. CONCLUSIONS

Much effort is being directed at building pervasive computing environments that provide more usable and compelling applications to the user. Agent-based computing is a relevant technology to implement this vision, because it captures many of the aspects used by pervasive computing applications, such as mentalistic attitudes, social behaviour and users' mobility. Agents will ultimately behave as humans by proxy in autonomic pervasive computing.

We believe that trust is a key notion in this paradigm and we are investigating ways to create trusted software agents. Our approach is based on several healthcare pervasive computing scenarios which were validated by subject matter experts. They form the first step in the definition of a realistic notion of trust. The approach incrementally uses scenarios' analysis in order to exhibit the important facets of trust which compose our trust analysis grid.

This grid encompasses many aspects previously discussed in the literature. Our approach provides a means to discover the relevant aspects out of a set of scenarios and relate them in a relevant way, though not systematicly. The first refinement of the grid presented here lead us to highlight three group of categories in the trust analysis grid: Subjective, System, and Data.

We are currently applying the trust analysis framework to other pervasive computing scenarios in order to further refine and stabilize our trust analysis grid. The healthcare pervasive scenarios will also be completed in order to take into account new categories of this grid.

We wish to explore how ontologies, the Semantic Web and Grids relate to these categories and may serve as a basis to define these facets of trust. The definition of trust that will stem from it will enable us to design a model. This model will be used in an agent-based system, with the ultimate goal to implement the pervasive computing scenarios.

9. ACKNOWLEDGMENTS

This work is supported by the Next Wave Technologies and Markets programme of the United Kingdom's Department of Trade and Industry.

10. REFERENCES

[1] AC Huang, BC Ling and S. Ponnekanti. Pervasive Computing - What is it Good For ? In *Proceedings of the ACM International Workshop on Data Engineering for Wireless and Mobile Access*, pages 84–91, Seatle, WA, USA, August 20 1999.

[2] Anthony D. Joseph (and other authors). Security, privacy, and health. *IEEE Pervasive Computing*, 2(1):96–97, January-March 2003.

[3] BBC News. Mobiles used to monitor asthma, March 3 2003. `http://news.bbc.co.uk/1/hi/technology/2808603.stm`.

[4] D. Harisson McKnight and Norman L. Chervany. The Meanings of Trust. Technical Report 94-04, Carlson School of Management, University of Minnesota, 1996. `http://misrc.umn.edu/wpaper/WorkingPapers/9604.pdf`.

[5] Diego Gambetta. Can We Trust Trust? In Diego Gambetta, editor, *Trust: Making and Breaking Cooperative Relations*, chapter 13, pages 213–237. Department of Sociology, University of Oxford, 2000. Electronic edition. `http://www.sociology.ox.ac.uk/papers/gambetta213-237.pdf`.

[6] Jennifer Golbeck, James Hendler, and Bijan Parsia. Trust Networks on the Semantic Web. In *Twelfth International World Wide Web Conference (WWW2003)*, Budapest, Hungary, May 20-24 2003.

[7] Josh Boyd. In Community We Trust: Online Security Communication at eBay. *Journal of Computer-Mediated Communication*, 7(3), April 2003. Electronic Journal, `http://www.ascusc.org/jcmc/vol7/issue3/boyd.html`.

[8] K. Ducatel, M. Bogdanowicz, F. Scapolo, J. Leijten, and J-C. Burgelman. Scenarios for Ambient Intelligence in 2010. Technical report, Information Society Technologies, European Commission, February 2001. `ftp://ftp.cordis.lu/pub/ist/docs/istagscenarios2010.pdf`.

[9] Lars Rasmusson and Sverker Jansson. Simulated Social Control for Secure Internet Commerce (position paper). In *Proceedings of the New Security Paradigm Workshop'96*, pages 18–26, Lake Arrowhead, CA, USA, September 16-19 1996.

[10] M. Ancona, G. Dodero, V. Gianuzzi, F. Minuto, and M. Guida. Mobile computing in a hospital: the "WARD-IN-HAND" project. In *Proceedings of the 2000 ACM Symposium on Applied Computing (SAC)*, volume 2, pages 554–556, Como, Italy, March 19-21 2000.

[11] Matt Blaze, Joan Feigenbaum, John Ioannidis, and Angelos D. Keromytis. The Role of Trust Management in Distributed Systems Security. In *Secure Internet Programming, Security Issues for Mobile and Distributed Objects*, LNCS 1603, pages 185–210. Springer-Verlag, 1999.

[12] Nicholas Hanssens, Ajay Kulkarni, Rattapoom Tuchinda, and Tyler Horton. Building Agent-Based Intelligent Workspaces. In *Proceedings of the International Conference on Internet Computing (IC'2002)*, volume 3, pages 675–681, Las Vegas, NV, USA, June 24-27 2002. CSREA Press.

[13] Paolo Bottoni, Maria Francesca Costabile, Stefano Levialdi, Maristella Matera, and Piero Mussio. Trusty Interaction in Visual Environments. In P. L. Emiliani and C. Stephanidis, editor, *Proceedings of the* 6^{th} *ERCIM Workshop "USER INTERFACES FOR ALL" (UI4ALL)*, pages 263–277, Florence, Italy, October 25-26 2000.

[14] Partha Dasgupta. Trust as a Commodity. In Diego Gambetta, editor, *Trust: Making and Breaking Cooperative Relations*, chapter 4, pages 49–72. Department of Sociology, University of Oxford, 2000. Electronic edition. `http://www.sociology.ox.ac.uk/papers/dasgupta49-72.pdf`.

[15] Pradip Lamsal. Understanding Trust and Security, 2001. `http://www.cs.helsinki.fi/u/lamsal/asgn/trust/UnderstandingTrustAndSecu%rity.pdf`.

[16] Raph Levien. Advogato's trust metric, 2000. `http://www.advogato.org/trust-metric.html`.

[17] Rino Falcone, Munindar P. Singh, and Yao-Hua Tan. *Trust in Cyber-societies, Integrating the Human and Artificial Perspectives*, volume 2246 of *Lecture Notes in Computer Science*. Springer, 2001.

[18] Ross J. Anderson. Information Technology in Medical Practice: Safety and Privacy Lessons from the United Kingdom, November 1998. `http://www.cl.cam.ac.uk/ftp/users/rja14/austmedjour.ps.gz`.

[19] Sarah Jones and Philip Morris. TRUST-EC: Requirements for Trust and Confidence in E-Commerce: Report of the Workshop held in Luxembourg, April 8th-9th. Technical Report EUR 18749 EN, European Communities EUR Report, 1999. Issue 2, `http://dsa-isis.jrc.it/TrustEC/D1.pdf`.