Assessing Interpersonal Trust in an Ambient Intelligence Negotiation System

Marco Gomes, Cesar Analide and Paulo Novais

Department of Informatics, University of Minho, Braga, Portugal
{marcogomes,cesar.analide,pjon}@di.uminho.pt

Abstract. This paper describes an approach to assess and measure trust based on a specific Ambient Intelligence environment. The primary aim of this work is to address and expand on this line of research by investigating the possibility of measuring trust based on quantifiable behavior. To do so, we present a brief review of the existing definitions of trust and define trust in the context of an Ambient Intelligence (AmI) scenario. Further, we propose a formal definition so that the analysis of trust in this kind of scenarios can be developed. Thus, it is suggested the use of Ambient Intelligence techniques that use a trust data model to collect and evaluate relevant information based on the assumption that observable trust between two entities (parties) results in certain typical behaviors. This will establish the foundation for the prediction of such aspects based on the analysis of people's interaction with technological environments, providing new potentially interesting trust assessment tools.

Keywords: Ambient Intelligence, Negotiation, Trust

1 Introduction

Trust is understood as a complex phenomenon and it is widely accepted as playing a significant role in human social relationships. Analyzing the phenomenon and its formal definitions we can find a profusion of interpretations containing different trust types or facets, with different properties. A situation in which requires different models for analysis. This statement is particularly evident when one go deeper into the primary disciplines concerned with trust relationships (psychology, sociology, etc.). This abundance of meanings inevitably leads to a degree of uncertainty about what is meant by trust, creating an enormous conceptual and terminological confusion. In fact, the most consistent thing about trust seems to be when it is absent [1]. Bearing in mind these diverging views as to what constitutes the appropriate definition of trust, the present work focuses on some specific definitions of trust, which are simplified so that these views can be captured algorithmically. Meanwhile, while a variety of definitions of the term *trust* have been suggested, in this work we will use a specific definition in which trust is seen as a rich and a complex mental attitude of x towards y as for a given action and goal [2]. Accordingly to Castelfranchi [2], this approach consists of evaluations of y and the situation, and of expectations about y's mind, behaviour and possible results. This includes, of course, the assumption that trust should be understood and can be interpreted by framing its natural subjectivity and the information needed at a particular time and for a specific context in a computer-based model: an abstraction that have the *power* to represent data in terms of entities and relationships relevant to a domain of inquiry (trust).

However, it is needed to find a way of evaluating (measure) trust to guide the research effort in the correct direction. Facing this issue, a critical question must be raised: how to balance the subjective nature of trust with the objectivity-dependent nature of a computer system? In other words, how can a computer system deal intelligently with this kind of subjectivity? Well, some glimpses of the ways that computer systems meet subjectivity can be found in the "expert" literature. This is exemplified in the work undertaken by Rosalind Picard in which she stressed that this endeavour is challenging but achievable [3]. Afterwards, she outlined a strategy for computer systems to cope with subjectivity issues. Specifically, it is proposed a three-fold approach: that they [computer systems] will need to (1) share some of the common sense of the user, (2) observe and model the user's actions, and (3) learn from these interactions. To summarize, this approach suggests that subjectivity is expressed as the user interacts with the system during a succession of queries. These involving inputs of the user can be tracked, modelled, and used to retrieve data consistent with changing requests. In our opinion, this strategy seems to be suitable to be applied so we planned to follow it in our work in ways to overcome the further issue. Hence, our motivation to examine trust is two-fold. First, the present study aims to address and expand on this line of research by investigating the possibility of measuring trust based on quantifiable behavior. To do so, we present a brief review of the existing definitions of trust and define trust in the context of an Ambient Intelligence (AmI) scenario. Further, we propose a formal definition so that the analysis of trust in this kind of scenarios can be developed. Thus, it is suggested the use of Ambient Intelligence techniques that use a trust data model to collect and evaluate relevant information based on the assumption that observable trust between two entities (parties) results in certain typical behaviors.

2 Measuring Trust using an Algorithmic Approach

As suggested in [4], trust literature can be categorized based on three criteria: (i) trust information collection, (ii) trust value assessment, and (iii) trust value dissemination. Each, in turn, can be further classified: trust information gathering from three sources, namely (i) attitudes, (ii) behaviors, and (iii) experiences; trust value assessment according to the data model, namely (i) graph, (ii) interaction, and (iii) hybrid; and trust value dissemination into trust-based recommendation and visualization models. About this work, and taking into account the Sherchan's strategy [4], we will highlight the trust definitions, trust types, properties, and measurement models from the perspective of the Computer Science and Economics disciplines. Such approach seeks to focus on some particular aspects of trust measurement, which are simplified so that these properties can be captured algorithmically. The simplifying process is attached to the need for obtaining and/or quantifies trust using detecting statistically significant trust-like behaviors. Therefore, the basis for our study is a proposition that trust results in characteristic interaction behavior patterns that are statistically different from random interaction in a computerized conflict and negotiation management system [5]. However, regarding the relationship between trust and negotiation one must stress that different research streams view trust in different ways depending upon the relationship under consideration [6]. So another important decision was to reduce the study of *trust domain* to the study of *interpersonal trust*, transforming the challenge of measuring trust slightly more accessible.

Facing these challenges, how trust measurements can be classified? Accordingly to [4], in general, it can be classified into two broad categories: "user" and "system". For the scope of this present work, we will consider only the notion of "user" trust is derived from Psychology and Sociology, with a standard definition as "a subjective expectation an entity has about another's future behavior" [4]. This implies that trust is inherently personalized. In this sense, trust is *relational*. As two individuals interact with each other frequently, their relationship strengthens, and trust evolves based on their experience. Following Adali proposition [7] related to interpersonal trust, the main type or facet of trust understudy will be the basis for our proposal. Hence, the focus will be on the following proposition: is possible to observe that trust between two entities *A* and *B* will result in certain typical behaviours that can be statistically captured. In other words, our aim is to quantitatively measure dyadic trust (trust between two entities) based on observed behaviour in a negotiation process. These behaviours are not only an expression of trust but can also facilitate the development of further trust. The simplest such behaviour is just interaction, in which an action occurs as two or more entities have some kind two-way effect upon one another. Regarding trust evaluation models, our approach to trust computation will be based on an interaction-based trust model. In this case, our interaction-based model evaluates trust based on the interactions performed in a computer-based conflict management system. Then trust-related information will be captured and assessed following the strategy aforementioned. The sources and the means to assess trust can be resumed to:

- **Source of trust information gathering**: Behaviours. Why? Because user behaviours are an important aspect of trust. Another reason is as they are identified by patterns of interactions they can be easily captured by an algorithm running on a computer system. Therefore, this will be our main source of information gathering. For example, in a negotiation scenario if a party is an extremely active participant and suddenly stops participating, this change in behaviour (he interaction is interrupted) is noticeable and might imply that this party's trust in the other party or with the party with whom he/she had been frequently interacting with has decreased.
- **Trust value assessment**: Regarding the techniques used to measure (compute) trust, they can be broadly classified into statistical and machine learning techniques, heuristics-based techniques, and behaviour-based techniques. In the present work, we will use behaviour-based techniques to assess trust. The reason for this choice lies in the fact that our measure of trust is based on quantifiable behaviour. So it seems obvious to choose assessment techniques based on this proposition.

After we have defined which directions and techniques we will embrace to measure trust, let us then formally define our proposal based on Adali's approach. This formalization is applied to the context of a negotiation, a process for two (or more) parties to find an acceptable solution to a conflict. Within a negotiation process, each party can make several proposals and exchange an unlimited number of messages. For the context of the interaction in a negotiation, the input is the proposal stream in a negotiation process, specified by a set of *4-tuples*,

$$\langle sender, receiver, message, time \rangle \tag{1}$$

note that we pretend to study the problem of trust purely from the observed interaction statistics, using no semantic information. Meanwhile, in our formalization, in an interaction between parties some semantic aspects (e.g. message) are considered in order to provide posterior semantic analysis. The output considered here is a set T induce from these inputs. The participants of the negotiation are represented by the elements of this set.

Despite what common sense stands in a *regular* situation (which more often interactions occur, the more likely that a trust relationship is likely to exist or to develop), the presence of distrust can imply a lack of interaction in conflict mitigation process. Assuming that, we postulate in this present work that the shorter and less balanced (means the average number of times that two entities interacts within the process) a interaction is between two parties, the more likely it is that they have a trust relationship; in addition, the more interaction there are between such a pair of elements, the less tightly connected they are. The fundamental task is first to identify when two elements of T set are interacting. Let A and B be a pair of users, and let $P = \{t_1, t_2, ..., t_k\}$ be a sorted list of the times when a message was exchanged between A and B. Therefore the average time between messages is defined as $\tau = (t_k - t_1)/k$.

The measure of trust will be based on the interactions in I, obeying the following postulates:(1)Longer interactions imply less trust;(2)More interactions imply less trust; and (3)Balanced participation by A and B implies less trust. We define the relational trust $R_I(A,B)$ as follows:

$$R_I(A,B) = \sum_{i=1}^{l} \|I_i\| \cdot H(I_i) \tag{2}$$

Where $H(I_i)$ is a measure of the balance in the i interaction contained in I. We use the entropy (measure of the amount of information that is missing in the flow of interaction) function to measure balance:

$$H(I_i) = -p \log p - (1-p) \log(1-p), \tag{3}$$

where $p(Ii)$ is the fraction of messages in the interaction Ii that were performed by A. The complexity of the algorithms for computing relational trust is $\Theta(|D| \log |D|)$, where $|D|$ is the size of the interaction stream.

3 Ambient Intelligence Applied to Trust Assessment

Ambient Intelligence has many uses in a wide domain. Under this paradigm, computational power is seamlessly embedded into the environment, ultimately creating computational environments that implement their life-cycle in an ideally invisible way for the user. Our primary aim is to develop AmI platform that will support the already electronic negotiation systems by providing relevant context information derived from the environment. In that sense, this work was decided to develop an environment that could be sensitive and responsive with both parties of a negotiation platform. As a result, this environment allows that all the components can be combined to implement complex functions. Namely, monitoring the negotiation process and perceiving how each issue or event is affecting each party, namely knowing the interpersonal trust of each party in each round.

Towards an intelligent conflict support system an ambient intelligence system was developed. The general working is to sense conflict context, acquire it and then make reasoning on the acquired context and thus acting in on the parties' behalf. To achieve this, the system builds up a profile and can link that profile subsequently with the correct individual performance within the conflict process that is monitored by the system. In other words, while the user conscientiously interacts with the system and takes his/her decisions and actions, a parallel and transparent process takes place in which contextual and behavioural information is sent in a synchronized way to the conflict support platform. The platform after converting the sensory information into useful data allows a contextualized analysis of the user's data. The contextualized analysis of user's data is critical when the data is from heterogeneous sources of diverse nature like sensors, user profile, and social media and also at different timestamps. To overcome some of these problems, the features are extracted from multiple sensor observations and combined into a single concatenated feature vector that is introduced into different classification modules (conflict styles, trust analysis, etc.). The multimodal evidence are integrated using a decision level strategy. Examples of decision level fusion methods employed in this work include weighted decision methods and machine-learning techniques and are detailed in previous work [8].

3.1 Case-Study: a Negotiation Scenario

A negotiation *scenario* should specify everything we know about the problem being addressed to frame the interactions that occur within. Taking this into account, a technological framework aimed to support the decision-making, by facilitating access to information such as the negotiation style of the parties or their social context, was adapted. In this work, it is introduced a new module that takes into account the context using trust analysis. Moreover, at this point, is highlighted that the primary objective of this research work is to identify and measure the users' interpersonal trust, to correlate to their negotiation performance and how it can be pointed out. Therefore, an experiment was set up in which we tried to estimate all the relevant aspects of the interaction between the individual that occur in a sensory rich environment (where contextual modalities were monitored). The participants of the proposed experiment were volunteers socially connected with our lab members. Twenty individuals participated, both female and male, aged between 19 and 42. The first step of the experiment was to ask the volunteers to fill in a small individual questionnaire. The next step was the monitoring of the individuals' interaction with the developed web-based negotiation game (in which subjects perform two distinctly different roles). During the experiments, the information about the user's context and performance (extracted using models based on behavioral and contextual monitoring [9], [10]) was provided through a monitoring framework, which is customized to collect and treat the interaction data. The participants played the web-based game through computers that allowed the analysis of the described features.

3.2 Results

After the experiment has been performed, the first step was to run some analysis to compare two sets of data under study: the data collected through the application of the questionnaire (measuring the participant's relationships) and the data gathered through the web-based negotiation game. According to our postulates, the basis for this analysis was the assumption that if A and B have a strong relationship then $R_I(A,B)$ value is below the median of the calculated trust for all the pairs within the interactions data set. In other words, if one pair of participants has a high degree of relationship then the same pair of participants, during the game, will perform fewer interactions than the median of total interactions per game. At this point, it should be highlighted that to apply non-parametric statistical analysis the raw data were pre-processed and it was subjected to tests. Therefore, the outcomes were compared using the Mann-Whitney U test (compares the central tendencies of two independent samples), given the fact that most of the distributions are not normal. The null hypothesis is thus: H_0 = The medians of the two distributions are different. For each two distributions compared, the test returns a p-value, with a small p-value suggesting that it is unlikely that H_0 is true. For each parameter (pair of participants), data from both samples is compared. In all the tests, a value of $\alpha = 0.05$ is used. Thus, for every Mann-Whitney test whose $p-$value $< \alpha$, the difference is considered to be statistically insignificant, i.e., H_0 is rejected. Consequently, the results have shown that no (statistically) important difference between data from the two samples were found. In other words, it means that our assumptions were valid. Another aim of this present work was to study the link in the relational trust between opponents and their behaviour exhibited in a negotiation scenario. So to analyse if trust relationships influence the negotiation performance, in the preliminary data analysis, the experimental data is organized into two groups based on the analysis of the trust measurements. One group contains the collection of some experimental data about how a user (A) behaves when he/she negotiates with someone (B) in which $R_I(A,B)$ have low value ($R_I(A,B) <$ median). In that sense, this approach will enable the establishment of a baseline for comparison with the second group, which comprises the data gathered from parties that negotiate with someone that haves high R_I values ($R_I \geq$ median). In particular, to statistically deal with data concerning the utility values of the parties' proposals, it was necessary to convert to an arbitrary numeric scale (0 is the least favorable style for the resolution and four the most favorable style). This type of scale means that the exact numeric quantity of a particular value has no significance beyond its ability to establish a ranking over a set of data points. Therefore, it was built rank-ordering (which describes order), but not relative size or degree of difference between the items measured. This was a mandatory step to make the data suitable for statistical and machine-learning techniques.

Moreover, the analysis shows that there is an apparent difference between the two groups regarding the negotiation styles exhibited during the game. One conclusion is that when participants share a significant trust relationship (low R_I value) the frequency of collaborative behaviours is far superior (49%) than otherwise (24%). In a similar analysis, but now concerning the roles played by participants, we conclude that the sellers are much more competitive than buyers (57% vs. 29%) while buyers are primarily collaborative. To interpret the significance of these results it is important to recall that participants were asked to negotiate a favorable deal in a competitive and win-lose scenario. Nevertheless, it is shown that when participants have a significant trust relationship they are more likely to transform it into a win/win situation. Something visible in the final results of the negotiations. On the one hand, we find that 100% of the agreements made by parties with a relevant trust relationship accomplished a successful deal, i.e., between the range of solutions that would benefit both. On the other hand, only 50% of negotiations that occurred between untrusted opponents (high R_I value) reached a mutual benefits agreement. It may be that they assumed they had to negotiate and get the best price (win/loose). But that was not the objective. Their objective was to negotiate a deal so they would not go bankrupt (win/win).

The preliminary evidence suggests a basis for expecting a connection between trust relationship and the use of negotiation styles. Despite these results, we still do not know much about how this kind of influence might facilitate (or inhibit) positive negotiation outcomes. Therefore, we will perform more and deeper experiments to understand how to collect and analysis relational ties that can influence negotiation performance.

4 Conclusion

We aim at firstly to identify and apply an algorithm for measuring interpersonal trust; secondly, to validate this approach opposing data collected from a questionnaire with data gathered from a web-based negotiation game to statistically study the correlation between mutual trust and conflict styles. From the experiment outcomes, the findings highlight the potentially quantifiable measurements of trust to further the understanding of negotiation dynamics. They pointed out relationships between the features being monitored and the participants' relationships elicited through a small questionnaire. These findings have the potential to enable the characterization of individuals and enhance negotiation performance. Thus the identification of trust relations between opponents in negotiation scenario in which can influence the negotiation performance is the main contribution of this work. Meanwhile, we can also conclude that these results are preliminary in the sense that there is more information that one can retrieve from the collected data namely through a more profound semantic analysis. This type of analysis could considerably enhance the trust measurements. Furthermore, due to the small sample size used in the current study, some caution must be taken when interpreting the results of the statistical analysis presented and underpinning the conclusions. In that sense, additional limitations of the current research must be pointed out. First, the participants were recruited from a particular population (that are socially related to our lab members)- a population that may limit the generalizability of the results. Admittedly, the participants of the experiment may not be representative of negotiation parties in general. Consequently, we are unable to demonstrate the causality of the variables conclusively. Moreover, it is possible that individual differences (i.e., personalities) might have influenced the impact of the results. Second, we tested all of the variables at the individual level that limited us from conducting global level analysis, which could provide us more variance of the data. This can be seen as another drawback of our study. Also, the data were collected through self-reported surveys at one time, which is subject to common method variance problem. Finally, the computational facet of this work should provide an understanding of the difficulties in algorithmically capturing and computing interpersonal trust in an AmI environment. A more comprehensive and in-depth study to provide theoretical advances, as well as implement technological solutions, is yet under development.

Acknowledgements

This work has been supported by COMPETE: POCI-01-0145-FEDER-007043 and FCT - Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within the Project Scope UID/CEC/00319/2013.

References

1. R.E. Fells. Developing trust in negotiation. *Employee Relations*, 15(1):33–45, 1993.
2. Cristiano Castelfranchi and Rino Falcone. Trust is much more than subjective probability: Mental components and sources of trust. In *Proceedings of the 33rd Hawaii International Conference on System Sciences-Volume 6 - Volume 6*, HICSS '00, pages 6008–, Washington, DC, USA, 2000. IEEE Computer Society.
3. Rosalind W. Picard. Computer learning of subjectivity. *ACM Comput. Surv.*, 27(4):621–623, December 1995.
4. Wanita Sherchan, Surya Nepal, and Cecile Paris. A survey of trust in social networks. *ACM Comput. Surv.*, 45(4):47:1–47:33, August 2013.
5. Davide Carneiro, Paulo Novais, Francisco Andrade, John Zeleznikow, and José Neves. Using case-based reasoning and principled negotiation to provide decision support for dispute resolution. *Knowledge and Information Systems*, 36(3):789–826, 2012.
6. William Ross and Jessica LaCroix. Multiple meanings of trust in negotiation theory and research: A literature review and integrative model. *International Journal of Conflict Management*, 7(4):314–360, 1996.
7. S. Adali, R. Escriva, M.K. Goldberg, M. Hayvanovych, M. Magdon-Ismail, B.K. Szymanski, W.A. Wallace, and G. Williams. Measuring behavioral trust in social networks. In *Intelligence and Security Informatics (ISI), 2010 IEEE International Conference on*, pages 150–152, May 2010.
8. Marco Gomes, Tiago Oliveira, Davide Carneiro, Paulo Novais, and Jos Neves. Studying the effects of stress on negotiation behavior. *Cybernetics and Systems*, 45(3):279–291, 2014.
9. André Pimenta, Davide Carneiro, Paulo Novais, and José Neves. *Hybrid Artificial Intelligent Systems: 8th International Conference, HAIS 2013, Salamanca, Spain, September 11-13, 2013. Proceedings*, chapter Monitoring Mental Fatigue through the Analysis of Keyboard and Mouse Interaction Patterns, pages 222–231. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.
10. Davide Carneiro, Marco Gomes, Paulo Novais, and José Neves. *Progress in Artificial Intelligence: 15th Portuguese Conference on Artificial Intelligence, EPIA 2011, Lisbon, Portugal, October 10-13, 2011. Proceedings*, chapter Developing Dynamic Conflict Resolution Models Based on the Interpretation of Personal Conflict Styles, pages 44–58. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.