2021 I-RIM Conference
October 8-10, Rome, Italy
ISBN: 9788894580525
DOI:10.5281/zenodo.6367943

A Framework for Trust-Based Task Assignment in Cooperative Teams of Robots

Alberto Grillo, Carmine T. Recchiuto, Antonio Sgorbissa
Laboratorium - DIBRIS, Università di Genova
Genova, Italy
albogrillo@gmail.com, carmine.recchiuto@dibris.unige.it, antonio.sgorbissa@unige.it

Stefano Carpin
School of Engineering, University of California Merced
Merced, USA
scarpin@ucmerced.edu

***Abstract*—In the last decade, the concept of trust and its dynamics has received considerable attention in robotics research. This is particularly true in the field of human-robot interaction, where several different factors, ranging from the user expectations of the robot capabilities to the physical appearance of the robot, has been defined as strongly affecting trust. On the contrary, the study of trust dynamics between robotic agents needs to be explored further. Starting from this premise, this work proposes a framework in which different robotic agents can model the concept of trust they have in each other for the accomplishment of a given task. Preliminary experiments performed with real robots (two Pepper robots and one NAO by SoftBank Robotics) provides a proof-of-concept for broader utilization of the system in cooperative robotic scenarios.**

***Index Terms*—Trust, Robot-Robot Interaction, Task Assignment, Social Robotics**

I. Introduction

The number of robots used in everyday activities is steadily increasing and is expected to keep growing. This will undoubtedly occur in industrial settings where the next generation of robots will be crucial in meeting the dynamic needs of collaborative and intelligent manufacturing that characterize the so-called *Industry 4.0 and Industrial Internet of Things* [1]. However, international market analyses anticipate widespread use of robots also by the general public [2]. In both industrial and domestic scenarios, different types of robots with different capabilities and marketed by different companies will need to work together, and possibly with humans, to achieve shared goals. Therefore, "teammates" will likely require to trust each other for efficient teamwork. In multi-agent systems, it is possible to define trust as "the subjective probability by which an agent (the trustor) expects that another agent (the trustee) succeeds in performing a given action on which its welfare depends" [3].

Robotic literature has primarily investigated trust in the context of human-robot interaction [4] [5], [6]. However, not all situations in which trust is involved require the trustee or the trustor to be human. Both in industrial and service scenarios, there are situations where autonomous robots must cooperate without having perfect knowledge about each other's capabilities. In this context, the trust that robots have in their own abilities versus other robots' capabilities may affect the selection of the most suitable candidate to perform the task.

In this work, we propose a novel framework for auction-based task assignment that can operate in a distributed open environment. The framework may handle heterogeneous robotic agents, with different perceptual, reasoning, and actuation capabilities, that are periodically assigned tasks to be performed to achieve a shared objective. The system takes inspiration from popular models in the literature, adapting them to an open environment in which robots may dynamically enter or exit, execute assigned tasks, or verify the correct execution of tasks by other robots. The purpose of such a framework is to allow agents to model trust in each other regarding the capability to accomplish an assigned task and use this model to evaluate possible candidates for task assignments whenever needed. In other words, agents may use the framework to outsource tasks they need to complete to achieve a particular goal.

II. Framework for Trust-Based Task Assignment

The general structure of the framework may be summarized as follows. First, using a portfolio of trust-related metrics, agents dynamically gather data about the other agents' capability to (i) perform actions; (ii) verify the outcomes of actions performed by other agents. Then, they will iteratively use and update these metrics to evaluate the trustworthiness of other agents, including themselves, during auctions, thus ultimately taking trustworthiness into account when taking a new decision for task assignment.

Given that, each agent in the framework should be able to:

- start a plan to manage one or more events;
- auction an action in its plan or bid on an action auctioned by another agent;
- execute or verify the execution of one or more actions;
- gather data from other agents about the success or failure of a given action;
- update its trust in other agents.

Concerning the latter, and considering a set of N_e events $\mathcal{E} = \{E_i\}$ (which may be triggered by an explicit command from a user, an alarm, or any other external stimulus that is processed by a robot and whose effect is a specific plan), a set of N_a actions $\mathcal{A} = \{A_i\}$, and a set of N_g agents $\mathcal{G} = \{G_i\}$, which can communicate with other agents, auction or bid for actions, we have defined the following context-independent metrics [5]:

- *Reliability*, one of the attributes of Logical Trust [4], which is the estimate, made by agent G_i, of the success rate of agent G_j in executing action A_k;
- *Verification Trustworthiness*, which measures the degree of the trustworthiness of a verifying agent depending on the consensus it has around its judgment skills. Formally, the *Verification Trustworthiness* measures how much the verification made by agent G_j about the success of action A_k is considered trustworthy according to an agent G_i and requires counting the number of opinions that are concordant or discordant with the judgment of G_j;
- *Perceived Competence*, which is computed, in its general formulation, as a function of the *Reliability* of G_j estimated by all agents $\{G_1 \dots G_n\}$ participating to the auctions as well as their own *Verification Trustworthiness*.

Regarding *Reliability*, it is worth noticing that it may be correctly computed only when a reasonably large sample of observations is available. To handle the transitory, i.e., when the agent G_j has not yet executed the action A_k a sufficiently large number of times, we have identified three possible strategies:

- a *Boot mode*, which requires the definition of a "boot phase" length, expressed as the number of A_k's auctions to which an agent G_j needs to participate before an auctioneer G_i starts using its *Reliability* as a measure of G_j's trustworthiness. In the boot phase, each auctioneer G_i shall trust G_j's declared *Reliability*, or rely on the outcomes of the other actions.
- a *Window mode*, similar to the *Boot mode*, but taking into account only a "memory window", instead of considering the full sequence of observations;
- a *BCI mode*, where each agent computes and shares not only the average *Reliability* of other agents, but also the binomial confidence interval (BCI) around such estimate that converges to zero as the number of executions increases.

Finally, the way in which individual *Reliability* and *Verification Trustworthiness* contribute to the *Perceived Competence* will depend on:

- the behaviour of agents towards the community: *individualistic* or *collectivistic*. *Individualistic* agents compute the *Perceived Competence* only based on their own observations; on the contrary, *collectivistic* agents take into account other agents' opinions by calculating a weighted average mediated by their *Verification Trustworthiness* (*Weighted Reliability*).
- the disposition of agents towards other agents: *optimistic*, *pessimistic*, *realistic*. The *optimistic/pessimistic/realistic disposition* of an auctioneer may play a key role to evaluate a bidder's trustworthiness. For example, in *BCI mode*, an *optimistic* auctioneer will consider the upper bound of the BCI to estimate the agent's *Perceived Competence*, therefore being more prone to forgive and give a second chance to an agent that failed the first attempts; a *pessimistic* agent will use the lower bound, thus being very conservative and cautious when it encounters a new bidder about which it has little data.

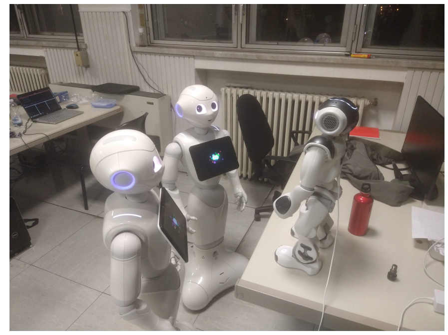

Fig. 1. NAO placed on the table to direct its speakers towards the two Peppers' microphones

III. Experiments

The framework has been tested in a real-world scenario comprising two SoftBank Pepper robots and one NAO robot. In order to simplify the execution of actions and their verifications, we have decided to associate actions A_1 and A_2 with sentences to be pronounced by the robots. Please notice that, while the robot that won the auction and pronounced the sentence always considers its action as a success, the other agents may judge outcomes differently. For instance, the distance between the speakers of a robot and the microphones of another robot, or the fact that NAO's and Pepper's microphones are placed over the robots' heads, may lead the speech recognition systems to fail.

Based on the considerations above, we performed experiments by placing NAO (agent G_3 in the following) on a table in front of the two Pepper robots (G_1 and G_2). NAO's speakers point towards Pepper microphones from above (Figure 1). The speech volume of the three robots is set to 80% of the maximum value. In this configuration, we expect that G_1 and G_2 can correctly verify the actions performed by all other robots, while G_3 will not be very performing in verifying the actions of G_1 and G_2 (given that its microphone is not oriented toward the two Pepper robots' speakers).

Robots work in *BCI mode*, with an *optimistic* disposition, and tests have been performed with robots showing both *individualistic* and *collectivistic* behaviours. Finally, please notice that event E_1 can only be handled by G_1, that will then auction A_1 (pronouncing the sentence "*Take the medicine*"); event E_2 can only be handled by G_2, that will then auction A_2 (pronouncing the sentence "*Drink some water*"). All agents can execute and observe both actions A_1 and A_2.

Results are shown in Figure 2, which reports, for each agent G_i how the estimated *Reliability* and *Verification Trustworthiness* of the three agents in the framework evolve as the number of auctions increase. In all plots, *Reliability* estimates are plotted with a continuous line, whereas *Verification*

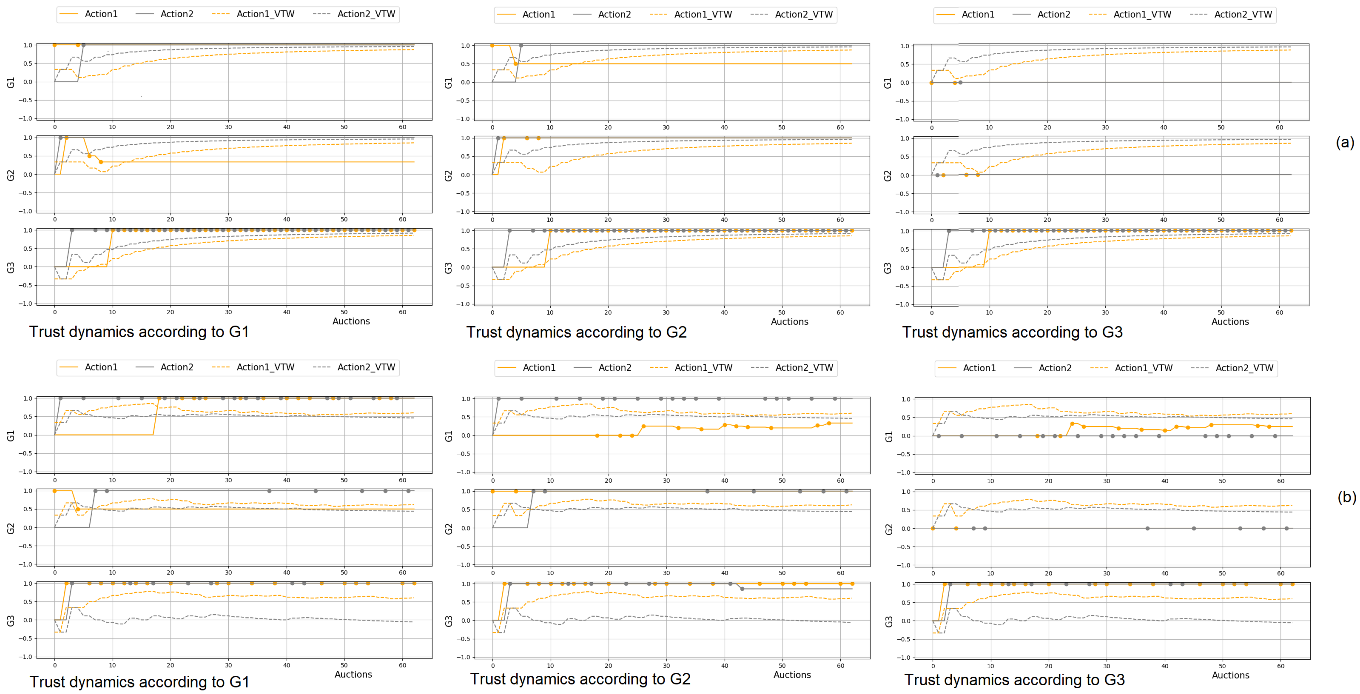

Fig. 2. Trust dynamics of the three agents with robots showing a *collectivistic* behaviour (a) and an *individualistic* behaviour (b).

Trustworthiness estimates are plotted with a dashed line. A small circle means that an action has been assigned to the corresponding agent at that time.

When robots exhibit a *collectivistic* behaviour (Figure 2 (a)) it can be observed that G_1 and G_2 can correctly recognize the actions performed by all other robots (even if in some cases the actions performed by G_2 are wrongly evaluated by G_1), while G_3 is not capable of correctly verifying the performance of G_1 and G_2. As a result, after a few iterations, the *collectivistic* auctioneers G_1 and G_2 tend to assign all actions to G_3, since it is considered the most reliable agent in terms of *Weighted Reliability* (i.e., it's the only agent whose actions are perceived as correctly executed by G_1, G_2, and G_3 itself).

When the robots show an *individualistic* behaviour (Figure 2 (b)), the actions are more uniformly distributed among all agents. Indeed, the two auctioneers G_1 and G_2 judge all agents equally trustworthy, and they only rely on their *Perceived Competence*. It may be also observed that, due to possible errors in speech-to-text conversion, the auctioneers may sometimes judge that an action has not been executed correctly. See for instance the trust metrics computed by G1: A_1 performed by G_2 during auction 4 is negatively evaluated by G_1 therefore producing a decrement in G_2's *Reliability*.

When this happens, the probability that the same action will be assigned again to a robot that has possibly failed decreases, since each auctioneer relies only on its own opinion. For the same reason, even if other agents judge that an auctioneer has failed an action, the auctioneer will ignore their opinions since all agents consider themselves as perfectly capable to execute actions. As a consequence, it can be observed in Figure 2 (b) that G_1 repeatedly assigns A_1 to itself even if it is considered very unreliable by G_2 and G_3.

Finally, it is worth noticing how the *Verification Trustworthiness* of the three robots is much lower in the second experiment. When robots are *collectivistic*, most of the auctions are won by G_3, whose actions are judged as correctly performed by all robots, but when robots are *individualistic*, actions are shared among all agents, and, concerning their verification, usually G_3 *disagrees* with G_1 and G_2 about the outcome of the actions.

These results, although preliminary, confirm the potentials and the functionalities of the presented framework, and its applicability to a real scenario.

References

[1] Gao, Z., Wanyama, T., Singh, I., Gadhrri, A., & Schmidt, R. (2020). From industry 4.0 to robotics 4.0-a conceptual framework for collaborative and intelligent robotic systems. Procedia Manufacturing, 46, 591-599.

[2] Yang, L., Henthorne, T. L., & George, B. (2020). Artificial intelligence and robotics technology in the hospitality industry: Current applications and future trends. Digital transformation in business and society, 211-228.

[3] Calvaresi, D., Najjar, A., Winikoff, M., & Främling, K. (2020). Explainable, Transparent Autonomous Agents and Multi-Agent Systems. Springer International Publishing.

[4] Cho, J. H., Chan, K., & Adali, S. (2015). A survey on trust modeling. ACM Computing Surveys (CSUR), 48(2), 1-40.

[5] Khavas, Z. R., Ahmadzadeh, S. R., & Robinette, P. (2020, November). Modeling trust in human-robot interaction: A survey. In International Conference on Social Robotics (pp. 529-541). Springer, Cham.

[6] Nahavandi, S. (2017). Trusted autonomy between humans and robots: Toward human-on-the-loop in robotics and autonomous systems. IEEE Systems, Man, and Cybernetics Magazine, 3(1), 10-17.