Towards a Social Robot as Interface for Tourism Recommendations

Berardina De Carolis[a], Pasquale Lops,[a], Musto Cataldo[a] and Giovanni Semeraro[a]

[a]*Department of Computer Science, University of Bari "Aldo Moro", Italy*

Abstract

The popularity of *social robots* is steadily increasing, mainly due to the interesting impact they have in several application domains. In this paper, we propose the use of Pepper Robot as an interface of a recommender system for *tourism.* In particular, we used the robot to interact with the users and to provide them with personalized recommendations about *hotels, restaurants, and points of interest in the area.* The personalization mechanism encoded in the social robot relies on soft biometrics traits automatically recognized by the robot, as age and gender, user interests and personal facets. All these data are used to feed a neural network that returns as output the most suitable recommendations for the target user. To evaluate the effectiveness of the interaction driven by a social robot, we carried out a user study whose goal was to evaluate: *(1)* how the robot affects the perceived accuracy of the recommendations; *(2)* how the user experience and the engagement vary by interacting with a social robot instead of a classic web application. Even if there is a large room for improvement, mainly due to the poor speech recognizer integrated in the Pepper, the results showed that the robot can strongly attract people, thanks to its presence and interaction capabilities. These findings encouraged us in performing a larger field study to test the approach in the wild and to understand whether it can increase the acceptance of recommendations in real environments.

Keywords

Social Robots, Recommender Systems, Tourism,

1. Introduction

Social robots are physically embodied, autonomous agents that communicate and interact with humans on a social and emotional level. They represent an emerging field of research focused on developing a "social intelligence" that aims to maintain the illusion of dealing with a human being [5]. Social robots are being applied in several domains such as assistance [17, 4], education [16], information providers in public places [13, 15], and they are are effective in influencing and motivating human behavior [14]. Indeed, it has been shown that users evaluate their experience more satisfying when interacting with humanoid robots and perceive robots as intelligent entities, establishing with them a relation similar to the one they have with other humans [12]. Thanks to their ability to enable a *very natural* interaction, social robots have shown a great potential in hospitality and tourism services [6] [8]. In this domain the exploitation of robots can lead to an improvement of the overall experience of the user [7] since the use of social robots allow to avoid cognitive overload and to reduce the time for finding information that are of interest.

However, web-based interfaces still represent today the most commonly used way to access to *tourism recommendations.* Accordingly, the current research was driven by a simple and straightforward research question: *how effective is the interaction driven by a social robot in a tourism recommendation scenario?*

In order to validate this claim, in this work we propose the use of the Pepper Robot as an interface of a tourism recommender system. In particular, we used the robot to interact with the users and to provide them with personalized recommendations about hotels, restaurants, and points of interest in the area. The recommendation strategy encoded in the social robot is organized in two steps: *(1)* user personal characteristics are gathered by using a soft biometrics module and a platform for holistic user modeling. The first automatically gathers user age and gender, while the seconds enriches the representation by including information about user preferences, mood, activities and personality traits; *(2)* given the representation of the user, a neural network previously trained with a huge set of users' stereotypes is run and a set of recommendations about hotels, restaurants and point-of-interests is returned to the target user.

Next, we carried out an experimental study aiming at understating whether the use of social robots can increase the perception of the quality of the recommendation and can improve the overall experience of the user. To this end, we compared the interaction with the social robot to a classic web-based interface. Our results support our hypothesis since the

CaESAR 2020: cAESAR 2020caesar2020.di.unito.it Workshop on Adapted intEraction with SociAl Robots
EMAIL: berardina.decarolis@uniba.it (B.D. Carolis)
ORCID:
© 2020 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-2724/paper8.pdf

robot provided the user with a satisfying experience. Indeed, attractiveness, stimulation and perspicuity (learnability) of the interaction were significantly better, dependability and efficiency were higher (but not significantly). Finally, the perceived accuracy and the overall satisfaction were also higher (not significantly) when Pepper was used.

To sum up, the contribution of this paper can be summarized as follows:

- We present a tourism recommender system based on neural networks, that identifies the characteristics of hotels, restaurants and point-of-interest that are suitable for a specific user;
- We integrate the recommender system in a Pepper social robot and we designed a voice-based and touchscreen-based interaction with the robot;
- We carried out a user study aiming at evaluating whether social robots can replace web-based interfaces in the task of providing users with recommendations for tourism.

The paper is organized as follows: in Section 2 provides a brief description of the Pepper4Tourism platform. Next, in Section 3 we present our experiment and the results of the study. Finally, Section 4 is aimed at a final discussion and a proposal for future developments.

2. Pepper4Tourism

The overall architecture of our tourism recommender system that exploits *Pepper social robot* is reported in Figure 1.

After a welcoming phase from Pepper, in which the purpose of the interaction is briefly explained, the user can interact with the robot to formulate her requests. In particular, user input can be provided via voice or through a touch screen. These modalities can be used indifferently. In the first case, the input is processed by the automated speech recognition module that is already encoded in the programming environment of the robot. In the second, the user can interact with the tablet of the Pepper robot showing the available choices and she can express her needs. It is worth to state that the touch screen is needed since the overall quality of the speech recognition module encoded in the Pepper is typically low. It is likely that by developing a more sophisticated speech recognition algorithm the usage of the tablet can be completely avoided.

In this implementation of the system, three different intents are caught: *hotels, restaurants and points of interest in the city*. Clearly, each intent invokes a different recommendation pipeline. To identify the intents, we adopted a a simple *keyword-based strategy* based on the presence of specific terms in the request, such as "hotel", "restaurant" or "point of interest" (or some synonyms, of course).

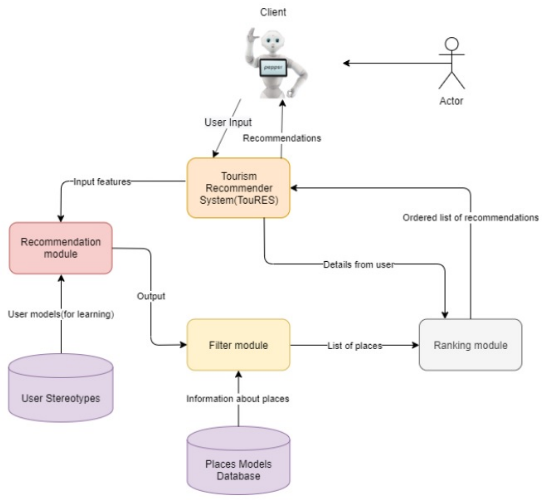

Figure 1: Architecture of the System

Soft Biometrics module. In order to start gathering information about the target user, a soft biometrics module is used. Soft biometric traits refer to physical and behavioral traits which are not unique to a specific subject, but are useful for identification, and description of human subjects[11].

In particular, the soft biometrics method implemented in Pepper allows to automatically infer *age* and *gender* of the user who is interacting with the robot. The algorithm follows the approach described in [10], that relies on a fine-tuned version of the VGG-16 architecture using unconstrained image dataset. Our approach showed good performance, since gender recognition has been performed with an accuracy of 85% while age estimation reached an accuracy (+/- 1 year) of 84% on the previously mentioned dataset. we tested the performance in real time in the wild having an accuracy of 87.5% for gender recognition and 62.5% for age estimation.

Recommendation module. Information about user gender and age are used to trigger the recommendation module implemented in the platform. To provide users with personalized recommendations we designed three different recommender systems (one for each *intent*) based on a feed-forward neural network with two hidden layers. The *input layer* of each network is devoted to the representation of the user and encodes 11 neurons, one for each feature describing the person.

In particular, as user features, we encoded age, gender, mood, physical activities, level of rest, overall physical state and five features that describe the Big-5 [19] characteristics of the user. Unfortunately, due to space reasons we can't go into details of our design choices and on the methodologies we exploited to obtain such features. In a nutshell, we can state that age and sex are obtained by using the previously described soft biometrics module, while the remaining features are gathered through a platform for Holistic User Modeling [18] that automatically infers users' characteristics by acquiring and processing social and personal data.

Next, the *hidden layers* of the networks encode 50 neurons. This is valid for all the intents. Such a value was set by evaluating different alternatives and by selecting the one having the lowest prediction error in a K-fold evaluation. Finally, the *output layer* encodes the available characteristics of the recommended items, as kind of restaurant, the services of the hotels and so on. For each neuron of the output layer the neural network can output 0 or 1. Clearly, 1 means that the recommended items should match a certain characteristic.

For each intent, to train the neural networks we manually built a set of 200 users stereotypes. Each stereotype maps the features of the user to the characteristics of the items they are (hopefully) more interested in. As an example, young users with a good mood and a high level of extroversion were associated to fast foods while older people that are tired and do not practice physical activities are mapped to a calm restaurant proposing fish or healthier foods.

Once the training of the network is completed, the system can provide recommendations. When a new user interacts with the recommendation module her features are obtained, the recommendation strategy is run and the characteristics that the recommendation should have is returned by the neural network. Such characteristics (e.g., calm place, cheap menu, etc.) are next used to filter and rank the available options. It is worth to note that our strategy is not affected by *cold start* problems. Indeed, thanks to the stereotypes we used as training, a new user that never interacted with the system can receive recommendations even at the first run.

Filtering module. Once the characteristics of the suitable items have been identified, the filtering module comes into play. This module has the goal of identifying the items that match the characteristics returned by the neural network. In a nutshell, the output of the neural network is a set of features $F = (f_1, f_2 \ldots f_n)$ the are supposed to be of interest for the user. Given these features, the filtering module filters out all the items that do not match not even one of the required characteristics (e.g., restaurants that are not fast foods).

Ranking module. Finally, the goal of the ranking module is to rank all the items that were returned by the filtering module in order to identify the *top-K recommendations*. In particular, the ranking is based on a linear combination of five different factors: *distance of the place from the current location of the user, popularity of the place, number of features returned by the neural network that are matched by the place, cost of the place*. By combining these factors a ranking of the items is obtained and the top-K recommendations are returned to the user.

The recommended items are finally shown on the tablet and the user may interact with them, get more information and express her feedback (see Figure 2).

3. Experimental Evaluation

To test our hypothesis, we carried out a user study aiming to evaluate: *(1)* the efficacy of a recommender system integrated in a social robot and *(2)* the overall experience of the user when the interaction was driven by the Pepper Robot instead of a typical web-based interface.

In particular, through this study we compared the use of Pepper to web-based platform and we evaluated the user experience in terms of *attractiveness, perspicuity, efficiency, controllability, stimulation, novelty*. These metrics are defined according to the standard User Experience (UEQ) questionnaire [9]. Moreover, we measured the time spent to get the recommendation in both the cases.

Participants. 52 people (30 males and 22 females) aged between 18 and 55 y.o, were recruited for the experiment. Most of them (70%) used regularly technologies such as smartphones and computers and the majority of them used TripAdvisor or Google for suggestions when travelling (90%).

Experimental Design. In order to validate our claims, we set up a scenario in which participants could ask for touristic information in our region. The 52 participants were divided into 2 groups equally distributed in number, gender and age. Half of the participants interacted with the robot and half with a web application. Both the group interacted with the same configuration of the recommendation module, of course.

In order to conduct the evaluation, we prepared a *pre-test* and a *post-test questionnaire*. The main purpose of the pre-test was to identify any possible bias in the results (e.g., the differences obtained in the post-

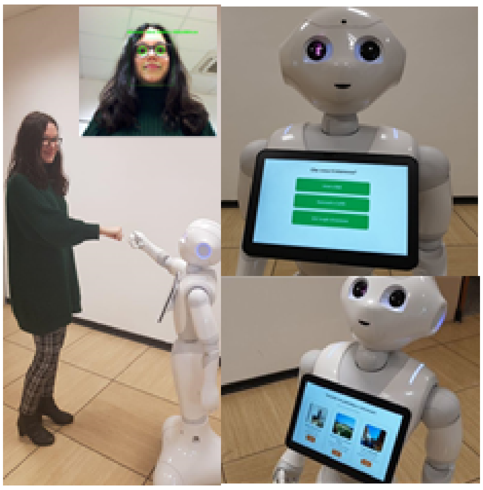

Figure 2: An example of interaction of a young woman with Pepper. In the upper part of the Figure the result of real time gender and age estimation is shown.

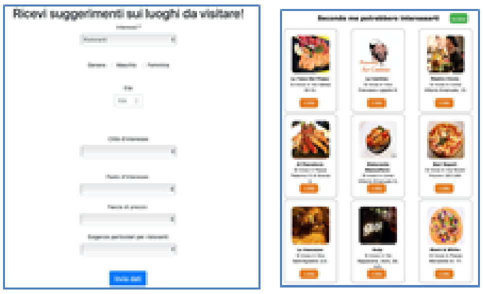

Figure 3: An example of the interface of the Web Application.

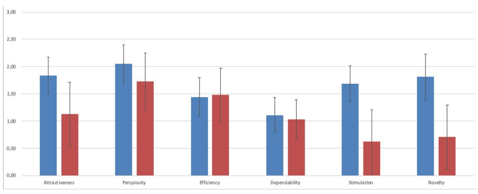

Figure 4: Result of the Experiment. The plot reports the scores obtained by the questionnaire for both the groups. Pepper = blue line, Web App = red line

test were due to difference in the use of technology and recommender systems.) Next, through the post-test - based on the previously mentioned UEQ questionnaire - we evaluated how the recommendation was perceived as effective and to what extent the overall interaction experience was positive.

Experimental Protocol. We started the experiment with a quick *training* phase, in which we aligned the users and we explained the purpose of the experiments. Next, we split the users and we asked them to fill in the questionnaire and to start interacting with Pepper (Group A) or with the Web Application (Group B).

Next, each participant was invited to interact with the recommender system (regardless the interaction mode) and to look for touristic information in the area, as places to visit, good restaurants or hotels. After the session, each of them was invited to fill out the post-test questionnaire.

In order to gather the features of the user, different strategies were adopted. In the Group A condition, *gender* and *age* were detected by using the soft biometrics module we previously mentioned. The recognition is performed by exploiting the RGB camera in the forehead with a resolution of 1280x960 and by detecting the face of a person entering in the field of view of the Pepper. It is worth to note that we also adapted the greetings by changing the level of formality/friendliness of the employed language according to the recognized age. As for the other features, we allowed the users to *(a)* sign-up to the Myrror platform, in order to automatically gather information from their holistic user profiles, or to *(b)* explicitly provide information about their mood, the amount of daily physical activities, level of rest and so on.

Conversely, the people in Group B had to interact with the web application thus they had to explicitly provide all the required information.

An example of the interaction with the Pepper robot is provided in Figure 2, while Figure 3 shows how the same information are presented on the web application accessed through a smartphone.

Results and Discussion. The analysis of the pre-test questionnaires confirmed our assumptions about the hypothetical user and ensured that differences in the post-test were not due to differences in the experience with technology. We do not report the results for the sake of brevity.

Next, we focused our analysis on the results of the post-test questionnaire in order to compare the user experience and the quality of interactions in both the two conditions. In particular, we used a *t-test* to compare the results provided by Group A with those provided by Group B in terms of Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation and Novelty.

As shown in Figure 4, the usage of Pepper improve the experience of the user for 5 out of 6 metrics, with the exception of the *efficiency*. This outcome was someway expected since users are more familiar with web-based interface than social robots. Conversely, for all the other metrics, the results obtained by Pepper are higher than those we obtained through the web applications. These differences are significant in terms of attractiveness, stimulation and novelty ($p<0.01$), and not significant for dependability and perspicuity.

Finally, as for the *quality of the recommendation*, the overall satisfaction concerning the received recommendation was higher for the Group A (*mean=3.3*) than for the Group B (*mean=2.3*). Statistical analysis showed a *significantly higher performance* ($t=2.716$; $p<0.02$) when Pepper was used instead of the web application.

4. Conclusions and Future Work

Social robots are being used in many domains and very often they are integrated in public spaces such as shops, airports, hotels and museums to attract the users and provide them with useful information. This work aims to provide a further validation of the role of social robots in providing the users with a satisfying experience with a tourism recommender systems.

In our opinion, a social robot such as Pepper is appropriate in this domain since it can be thought as an interactive InfoPoint which is more attractive than the classic displays that are typically used. In particular, in this preliminary experiment we tried to demonstrate that the interaction with social robots is preferred by humans than web-based one and that also the overall satisfaction with the received recommendation is higher when it is provided by the Pepper robot.

The results of the study showed the potential of social robots in this domain, since they emerged as a more engaging and effective interface for such a recommender system. These findings encourage us to perform a bigger user study in order to better understand the effects of social robots on the perceived quality of the recommendation.

In particular, we plan to investigate how the different characteristics of the users (e.g., personality traits) or the different groups (based on age or on the familiarity with these technologies) impact on the overall perception of the social robots in the area of tourism recommendation. Moreover, we will also analyze different algorithms in terms of design of the neural network, by evaluating different (and more sophisticated) architectures.

References

[1] Schwartz, Barry. The paradox of choice: Why more is less. 2004, HarperCollins

[2] Staffa M., Rossi S. Recommender Interfaces: The More Human-Like, the More Humans Like. Social Robotics. ICSR 2016. Lecture Notes in Computer Science, vol 9979. Springer, Cham. 2016, pp.94–103

[3] Ivanov, Stanislav and Gretzel, Ulrike and Berezina, Katerina and Sigala, Marianna and Webster, Craig. Progress on robotics in hospitality and tourism: a review of the literature. doi:10.1108/JHTT-08-2018-0087. 2019

[4] Joost Broekens and Marcel Heerink and Henk Rosendal, Assistive social robots in elderly care: a review, Gerontechnology, 2009, pp.94–103

[5] Alqaderi, Mohammad and Rad, Ahmad 2018 A Multi-Modal Person Recognition System for Social Robots Applied Sciences, doi:10.3390/app8030387

[6] Tung, Vincent and Law, Rob 2017 The potential for tourism and hospitality experience research in human–robot interactions, (29). International Journal of Contemporary Hospitality Management, doi:10.1108/IJCHM-09-2016-0520

[7] Santos, Joana and Campos, Daniel and Duarte, Fábio and Pereira, Filipe and Domingues, Inês and Leão, João and Xavier, José and de Matos, Luís and Camarneiro, Manuel and Penas, Marcelo and others. A Personal Robot as an Improvement to the Customers' In-Store Experience. Service Robots. 2018

[8] Ivanov, Stanislav and Gretzel, Ulrike and Berezina, Katerina and Sigala, Marianna and Webster, Craig. Progress on robotics in hospitality and tourism: a review of the literature. 2019. doi:10.1108/JHTT-08-2018-0087

[9] Schrepp, Martin. User Experience Questionnaire Handbook. 2015

[10] Rothe, Rasmus and Timofte, Radu and Van Gool, Luc. Dex: Deep expectation of apparent age from a single image.Proceedings of the IEEE International Conference on Computer Vision Workshops. 2015

[11] Dantcheva, Antitza and Velardo, Carmelo and D'angelo, Angela and Dugelay, Jean-Luc. Bag of soft biometrics for person identification. Multimedia Tools and Applications. v. 51 n. 2 pp. 739-777. 2011.Springer

[12] Bartneck, Christoph and Bleeker, Timo and Bun, Jeroen and Fens, Pepijn and Riet, Lynyrd. The influence of robot anthropomorphism on the feelings of embarrassment when interacting with robots. Paladyn - Journal of Behavioral Robotics. V 1. n. 2, pp. 109-115 doi=10.2478/s13230-010-0011-3. 2010

[13] Iolanda Leite and Carlos Martinho and Ana Paiva. Social Robots for Long-Term Interac-

tion: A Survey. International Journal of Social Robotics. vol 5. pp. 291-308 2013.

[14] Fasola, Juan and Mataric, Maja. Using Socially Assistive Human–Robot Interaction to Motivate Physical Exercise for Older Adults.Proceedings of the IEEE. 2012. doi: 10.1109/JPROC.2012.2200539

[15] Acharya, Urja and Bevins, Alisha and Duncan, Brittany A. Investigation of human-robot comfort with a small unmanned aerial vehicle compared to a ground robot. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2017

[16] O Mubin and CJ Stevens and S Shadid and Al Mahmud A and JJ Dong. A review of the applicability of robots in education. Technology for Education and Learning. 2013 doi:10.2316/Journal.209.2013.1.209-0015

[17] Palestra, Giuseppe and Varni, Giovanna and Chetouani, Mohamed and Esposito, Floriana. A multimodal and multilevel system for robotics treatment of autism in children. 2016. doi:10.1145/3005338.3005341

[18] Musto, Cataldo and Semeraro, Giovanni and Lovascio, Cosimo and de Gemmis, Marco and Lops, Pasquale. Myrror: a Platform for Quantified Self and Holistic User Modeling. Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization pp.215–216 2018

[19] Goldberg, Lewis R. The structure of phenotypic personality traits. American psychologist (48) 1993. American Psychological Association