Modeling and Understanding Future Action Decisions of Players during Online Gaming

Fabrizia Auletta
Macquarie University
Sydney, Australia
fabrizia.auletta@hdr.mq.edu.au

Gaurav Patil
Macquarie University
Sydney, Australia
gaurav.patil@mq.edu.au

Rachel W. Kallen
Macquarie University
Sydney, Australia
rachel.kallen@mq.edu.au

Mario di Bernardo
University of Naples Federico II
Naples, Italy
mario.dibernardo@unina.it

Michael J. Richardson
Macquarie University
Sydney, Australia
michael.j.richardson@mq.edu.au

ABSTRACT

Contemporary Supervised Machine Learning (SML) and explainable AI (artificial intelligence) methods can be employed to both model and understand the decision making behavior of human actors within a multi-agent task setting. Here, we apply such modeling approach to capture the decision-making behavior of human actors playing a 3-player online herding game called "Desert Herding". Of particular interest is whether the modeling approach can be employed to predict and understand the target switching strategies of human herders at variable prediction horizons and whether the explainable AI tool SHAP can be leveraged to identify the key informational variables (features) underlying the players' target selection decisions.

KEYWORDS

decision-making, supervised machine learning, explainable-AI, artificial neural networks, joint-action, multi-agent interaction

ACM Reference Format:

Fabrizia Auletta, Gaurav Patil, Rachel W. Kallen, Mario di Bernardo, and Michael J. Richardson. 2022. Modeling and Understanding Future Action Decisions of Players during Online Gaming. In *Proceedings of the 10th International Conference on Human-Agent Interaction (HAI '22), December 5–8, 2022, Christchurch, New Zealand.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3527188.3563926

1 INTRODUCTION

Pivotal to the structural organization of multiplayer behavior is the ability of co-actors to effectively decide how and when to act [1, 10, 11]. Understanding the decision making processes that lead to effective multiplayer behavior is therefore fundamental for developing an understanding of task- and team- working skills [8, 12], as well as human decision-making in general. Moreover, due to the rapid advances in interactive robotic or artificial agents, and their increasing presence within our social world, understanding and modeling human decision making during multiplayer (and multiagent) behavior also has significant implications for the development of such technologies [4, 5].

HAI '22, December 5–8, 2022, Christchurch, New Zealand

ACM ISBN 978-1-4503-9323-2/22/12…$15.00
https://doi.org/10.1145/3527188.3563926

2 MULTI-PLAYER DECISIONS

Recent research has demonstrated that the decision making behavior of human actors can potentially be modeled using supervised machine learning – human task data is used to train artificial neural networks to predict human task action decisions – at timescales well in advance of a human actor's conscious (or unconscious) intent [2, 3]. Furthermore, that the explainable-AI technique SHapley Additive exPlanation (SHAP, [7]) can then be used to uncover what task information supports effective actors decisions [3].

To explore this possibility further, we extracted and modeled the target selection decisions of players who completed a online, 3-player video game called "Desert Herding" as part of a study conducted by Prants *et al.* [9]. Such a game is representative of the emerging interactions in teams who need to coordinate to successfully achieve a shared goal, in this case, to contain wandering target agents. Most in general, the "desert herding" is also a paradigm of multi-agent interactions as, e.g., evacuation scenarios or, more trivially, field trips where teachers need to lead students. For this game, player teams were required to retrieve and corral autonomous target agents (small robots) randomly spread around a large desert game area into a specified containment region. Ten three-person teams completed four separate 1-hour game sessions, in which the number of targets (herd size), visibility and the presence (or not) of an Heads-Up Display (HUD) was manipulated; see [9] for a complete description of the task and task manipulations. Here, we focused on modeling the individuals decision-making process during the final experimental session for both the high (i.e., >150m visibility horizon) and low (i.e., a 15 m visibility horizon) visibility conditions, when the participants were tasked to corral $N_T = 9$ targets, and the HUD was present. Note that the HUD would always show the position of all $N_H = 3$ herder and $N_T = 9$ target agents w.r.t. cardinal points as in Figure 1. We chose Session 4, as by this final session all teams had reached a high level of performance. A total of 20 successful trials were obtained resulting in a 20 trial x 3 player data set.

To investigate the decision making processes of team members we selected as input to a Long-Short-Term-Memory neural network

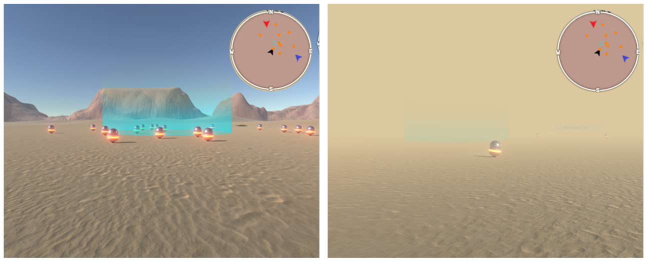

Figure 1: Examples of the multiplayer desert herding environment, first introduced in [9], under high (left panel) and low (right panel) visibility conditions.

(LSTM$_{NN}$), N_{train} time series of N_{sv} state variables, of length $t_f - t_i = T_{seq}$, with the output label corresponding to the ID of the target agent being corralled at $t_{f+\tau_{hor}}$, with $\tau_{hor} > 0$. The output label (i.e., the ID of the target that a player was currently corralling), was automatically determined from the recorded data. At each time step, the target a player was corralling was identified as the targeted agent (already numbered 1 to 9 in the data recordings), such that the prediction output was set to ID = 0 if a player was not currently engaged with (corralling) a target or ID = $1, 2, \ldots, N_T = 9$, otherwise. We chose a sampling time $dt = 0.2$ s, a sequence length $T_{seq} = 5$ s and a variable prediction horizon τ_{hor} motivated by the players inter-target movement times (i.e, decision making timescale) or, computationally, as the time it took a player to move outside the repulsive radius of a target and into the repulsive radius of the following one (Figure 2).

As illustrated in Figure 2, the inter-target decision-making timescales of the human players was distributed between 0.07 s and 11 s, indicating that in order to fully understand the decision making behavior of the players we indeed needed to model targeting switching events; that is, the variable timescale decision events that corresponded to when participants decide to change which target agent to corral. To this end, we trained LSTM$_{NN}$s to predict the ID of the next engaged target, independently from how far in the future this target switching event was observed. More specifically, for each sequence of the system evolution, the output label was the ID of the next target corralled that was different from the current target engaged with at t_f, with one limiting restriction, $\tau_{hor,max}$, which was the maximum step in the future explored.

We considered a minimum and maximum prediction horizon $t_{\tau_{hor,min}} = dt$ and $t_{\tau_{hor,max}} = 20$ s (or 100 times steps in the future), resulting in 97.4% switching samples for the full-visibility condition and 97.3% for the partial visibility condition. Note that switching samples represent the cases that either a participant started engaging a target (ID = 0 to ID $\neq$ 0), or a participant stopped engaging a target (ID $\neq$ 0 to ID = 0) or that it switched between engaged targets.

Prediction performance, averaged over 100 sets of $N_{test} = 2000$ samples, confirmed that both the partial and full visibility LSTM$_{NN}$s models were able to predict the next corralled target with an accuracy above 91% $\pm$ 0.5. A Kruskal Wallis statistical test [6] on prediction performance metrics indicated that the difference in the accuracy of the full and partially visibility models was statistically

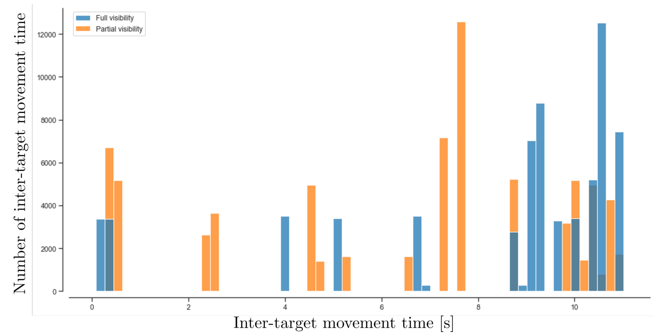

Figure 2: Distribution of inter-target movement times [s] averaged across player for the 20 successful trials employed for model training and testing.

significant ($\chi^2 = 80.75$, $p < 0.0001$), with the full visibility model slight outperforming the partial visibility model (see Table 1 for more model performance metrics). Given the high model perfor-

	Accuracy	Precision	Recall
Full visibility	92.15 ± 0.54	92.39 ± 0.54	93.45 ± 0.51
Partial visibility	91.25 ± 0.61	91.35 ± 0.63	92.14 ± 0.59

Table 1: Average performance [%] of the multi-label predictor trained on time-series of length $T_{seq} = 5$ s, and variable prediction horizon.

mance, we then employed SHAP [7] to better understand how the different input features influenced model predictions. Note that the SHAP algorithm pairs each input feature with a SHAP value representing the weight that feature has on the model output; the higher the SHAP value, the greater the influence that feature has on the model output. Global feature importance can therefore be assessed by calculating the averaged SHAP value for a given feature over the test set.

For each LSTM$_{NN}$ trained to predict target switching decisions at a variable prediction horizon, we computed the SHAP values for $N_{test} = 6000$ samples and rank-ordered each feature in terms of its average importance (i.e., average SHAP value).

Figure 3 provides visual illustrations of feature importance for the partial and full visibility models, with the top twenty most prevalent features for non-zero prediction outputs (i.e., ID = 1 to 9). As can be discerned from an inspection of this Figure, how far (or close) players were from target agents (i.e., herder-to-target distance) always ranked as top features across target ID – was always within the top twenty most important features for both the partial and full visibility models. Independent of visibility condition, a player velocity and target acceleration were also key to target predictions. A player's (filled blue square mark) velocity was key to target selection decisions for 100% of all target ID predictions, with the to-be-corralled targets acceleration (filled orange circle marks) important for over 90% of prediction outcomes. The importance of these features was likely due to (i) players selecting new targets just after they had corralled a target into the containment area, with the to-be-corralled target typically stationary (i.e., had zero or close to zero velocity) and (ii) players choosing targets that were not increasing or decrease speed, as the later only occurred if a target was already been corralled by another player.

With regard to what differentiated the player target selection decisions in the partial and full visibility conditions, the distance of the targets from the containment goal area was weighted as more important (on average) for partial visibility model predictions compared to the full visibility model. One possible reason being that in the low visibility condition players could often not see any target within their first person field of view and, hence, relied more heavily on the HUD to select targets, naturally choosing targets that were furthest from the containment area.

3 CONCLUSIONS

In conclusion, the current study extended the SML and explainable AI pipeline introduced in [2, 3], demonstrating how the model approach can be used to predict and understand the target selection decisions of human actors during a multiplayer (3-person) online game. Current research is exploring whether these models can also be used to create 'human-like' interactive artificial agents for robust human-machine teaming and training systems.

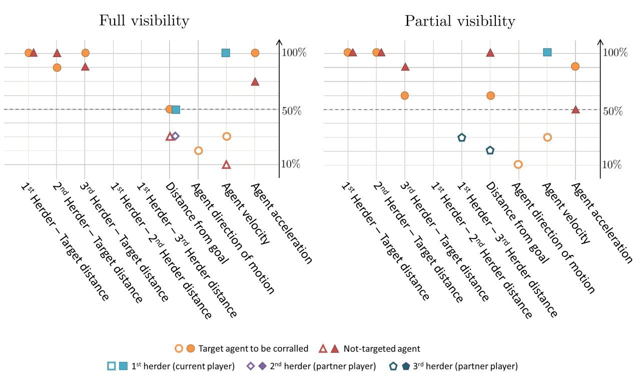

Figure 3: Explanation of the Multi-label predictor of "agent to be targeted" prediction outputs, for full and partial visibility conditions. Feature type is listed on the x-axis, with the y-axis representing the portion of $N_T = 9$ targets for which a state variable was found on average to be a top twenty global feature. Filled marks indicate state variables used to predict a portion of targets equal or above 50% while empty marks indicate state variables used to predict smaller portions.

ACKNOWLEDGMENTS

This project was supported by an Industrial and International Leverage Award from University of Bristol, an International University Research Excellence Scholarship Scheme from Macquarie University, an Australian Research Council Future Fellowship (FT180100447), and the Australian Department of Defense, Science and Technology (DST) group (partnership grant MyIP8655).

REFERENCES

[1] Duarte Araujo, Keith Davids, and Robert Hristovski. 2006. The ecological dynamics of decision making in sport. *Psychology of sport and exercise* 7, 6 (2006), 653–676.

[2] Fabrizia Auletta, Mario di Bernardo, and Michael J. Richardson. 2021. Human-inspired strategies to solve complex joint tasks in multi agent systems. *IFAC-PapersOnLine* 54, 17 (2021), 105–110. https://doi.org/10.1016/j.ifacol.2021.11.033

[3] Fabrizia Auletta, Rachel W Kallen, Mario di Bernardo, and Micheal J Richardson. 2022. Predicting and Understanding Human Action Decisions during Skillful Joint-Action via Machine Learning and Explainable-AI. *arXiv preprint arXiv:2206.02739* (2022).

[4] Wolfram Erlhagen and Estela Bicho. 2006. The dynamic neural field approach to cognitive robotics. *Journal of neural engineering* 3, 3 (2006), R36.

[5] Xiaocong Fan, Michael McNeese, and John Yen. 2010. NDM-based cognitive agents for supporting decision-making teams. *Human–Computer Interaction* 25, 3 (2010), 195–234.

[6] William H Kruskal and W Allen Wallis. 1952. Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association* 47, 260 (1952), 583–621.

[7] Scott M Lundberg and Su-In Lee. 2017. A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems* 30 (2017), 4765–4774.

[8] Judith H Martens. 2020. Habit and Skill in the Domain of Joint Action. *Topoi* 1 (2020), 1–13.

[9] Matthew James Prants, James Simpson, Patrick Nalepka, Rachel W Kallen, Mark Dras, Erik D Reichle, Simon Hosking, Christopher J Best, and Michael J Richardson. 2021. The structure of team search behaviors with varying access to information. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, Vol. 43.

[10] Michael J Richardson, Kerry L Marsh, and Reuben M Baron. 2007. Judging and actualizing intrapersonal and interpersonal affordances. *Journal of experimental psychology: Human Perception and Performance* 33, 4 (2007), 845.

[11] Natalie Sebanz, Harold Bekkering, and Günther Knoblich. 2006. Joint action: bodies and minds moving together. *Trends in cognitive sciences* 10, 2 (2006), 70–76.

[12] N Sebanz and G Knoblich. 2009. Prediction in joint action: what, when, and where. *Topics in Cognitive Science* 1, 2 (2009), 353–367.