A Human-Aware Decision Making System for Human-Robot Teams

Saurav Singh
Electrical and Computer Engineering Department
Rochester Institute of Technology
Rochester, NY, USA
ss3337@rit.edu

Jamison Heard
Electrical Engineering Department
Rochester Institute of Technology
Rochester, NY, USA
jrheee@rit.edu

***Abstract*—In a human-robot team, robots can perceive the surrounding environment states using various sensors, but typically do not perceive the humans' internal states (e.g., workload, fatigue, and comfort). This work presents a human-aware system of systems that incorporates human states into the robot's decision-making process to achieve more fluid human-robot team dynamics and improve the overall team performance. The human-aware system architecture employed a reinforcement learning paradigm as the decision-making agent that incorporated task and human state information. The architecture was validated on teaming data from the NASA MATB-II task environment. The results suggest that the learned action strategies were fine-tuned to the human.**

***Index Terms*—Soft Actor Critic; Workload; Human Aware System; Reinforcement Learning**

I. Introduction

An effective human-robot team requires a robust understanding of the team goals, the environment, and the teammate's states. Robots perceive their physical surroundings using various sensors, such as LiDAR, depth, thermal, Radar, ultrasonic, and cameras. However, the robot typically does not consider the human teammates' internal states, such as workload, and fatigue. These states have a direct impact on human performance; thus, influencing the overall team performance. This can also impact robot performance indirectly as the robots' actions are largely influenced by humans in different forms of human-robot interaction such as teleoperation, supervised and reprogrammed by humans, or as teammates in human-robot collaboration [1], further affecting the overall team performance. Team performance is especially critical in high-stress multitasking environments (e.g., search & rescue operations), where sub-optimal team performance may lead to high failure costs. Physiological computing systems address this gap by taking the human states and/or physiological data into account to adapt their behavior in order to improve the overall team performance [2].

Human states are typically estimated subjectively using questionnaires [3]. A lot of studies have also estimated human states from objective data, usually physiological (e.g., heart rate, heart rate variability, respiration rate, electromyography, galvanic skin response, speech features, noise level, posture, electroencephalogram, and eye-tracking [4] [5]).

Human workload is used as the human state for validating the human-aware system architecture presented in this work. Critical high-stress multitasking environments with high failure cost such as NASA control room, piloting an airplane, search & rescue operations require humans to deliver optimal performance, considerably increasing the human workload. This can decrease human performance due to an overload workload state [6]. If human is underloaded, they may fail to take any needed actions timely as they can become disengaged from the system. The overall human workload is composed of cognitive, auditory, speech, visual, fine motor, tactile, and gross motor. The fine motor, tactile, and gross motor components can be combined into physical workload [7].

This paper seeks to improve overall human-robot team performance using a human-aware system of systems paradigm that relies on reinforcement learning to model the various trends and relationships between task states, human internal states, and overall team performance. This paradigm was evaluated during a human subjects study using the NASA Multi-Attribute Task Battery (MATB-II) [8], which is a supervisory human-robot teaming system.

The key contributions are:

- A human-aware decision making system that determines appropriate robot adaptations using task and human workload information.
- A reinforcement learning-based approach that can learn trends between human states and team performance.

The rest of the paper is organized as follows: Section II presents related work in the field of physiological computing systems and adaptive automation. Section III presents the human-aware system architecture and details its different subsystems in contrast to related work. Section IV describes the human-subjects experiment used to validate the architecture. Section V presents the experimental results and discusses the findings. Section VI concludes this paper.

II. Related Work

Human-Robot teaming research has typically focused on human external states (e.g., position, velocity, head pose, arm pose, and gaze), but recent efforts have incorporated human internal states (e.g., emotions, workload, fatigue, and stress) [9]. Human internal states can be subjectively estimated using questionnaires (e.g., NASA Task Load Index for workload [3]). However, such questionnaires record the human's perceived

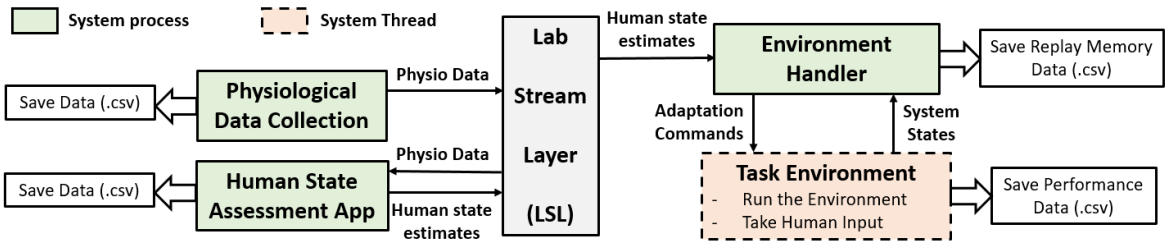

Figure 1. Human Aware system architecture for Human-Robot Teaming.

state in an intrusive manner and can be biased, limiting real-world use. Changes in the human states can also be captured using physiological data. Many studies have proposed methods to estimate various human internal states using objective measures (eg., physiological data), which makes the use of human states feasible in real-time automation systems.

Liu et al. [10] presented an EMG patch to estimate muscle fatigue during exercise, which can potentially be used in a human-robot collaboration system to accommodate for reduced human physical capabilities due to muscle fatigue. M. Ding and others [11] used a combination of EMG and joint torque signals during the lift up by a nursing care robot to estimate human comfort, however, it was a human factors study with no adaptations based on the estimated comfort. Rodrigues et al. [12] developed a stress estimation model using ECG, EMG, and galvanic response features, which can be used to assess human stress in high-stress environments. These estimates have great potential to deliver information about human sub-optimal states that can affect human performance, and this work presents a method to effectively use these human state estimates for robot automation decisions.

This work uses human workload to validate the presented approach. Many researchers have proposed methods to estimate human workload subjectively using questionnaires (e.g., the NASA-TLX [3]), or objectively physiological metrics [13] or a combination of both using machine-learning algorithms [14]. Human workload has been used to adapt and improve human-robot team performance with rule-based approaches while using only a subset of human's workload state (i.e., cognitive workload) [15] [16]. The approach presented in this study uses all the components of the human workload state [17] and compares the reinforcement learning approach with traditional rule-based adaptation schemes [18].

III. Human aware System architecture

A human-aware system of systems must have a mechanism to estimate and incorporate human internal states into its decision-making process. Figure 1 shows the developed human-aware system of systems architecture, which incorporates a physiological data collection system, a human state assessment system, and an adaptive environment handler system,

A. Physiological Data Collection System

The Physiological Data Collection system contains a sensor suite comprising a Zephyr BioHarness (a chest harness device used to collect heart rate, heart rate variability, respiration rate, and posture), a decibel meter (measures surrounding noise level), and a microphone. The system then sends processed heart rate, heart-rate variability, respiration rate, posture, noise level, speech pitch and intensity, and syllables per second into the Lab Stream Layer (LSL) [19]. LSL is a system that handles the networking, time-synchronization, and near real-time collection and access to collected time series data (e.g., physiological data). The processed physiological data is then used within the human state assessment system.

B. Human State Assessment App

The human state assessment system extracts relevant features from the incoming and previous physiological data. The resulting features are then used to detect the human's internal state. This work adapts a prior workload assessment algorithm [4] to estimate the human's workload state using the mean, standard deviation, average gradient, and slope of the physiological data with a 30-second window and 5-second stride. The features are then fed into a neural network consisting of three fully connected layers with rectified linear unit activation functions and 64, 64, and 16 neurons in the respective layers. The output layer contained 5 neurons with a linear activation function representing the cognitive, physical, speech, visual, and auditory workload estimates. An overall workload estimate is computed by summing each component estimate. The modified algorithm was validated on previously collected data prior to incorporation into the system. The resulting workload estimates are then sent back into the LSL for use in the task environment handling system.

C. Task Environment Handler

An environment handler was used to handle the automation decisions for the task environment. It was responsible for accessing the human state estimates from the LSL layer, determining environmental states, and using this information to make automation decisions. These decisions were then communicated to the task environment.

1) Task Environment: NASA MATB-II task environment [8] depicted in Figure 2 was used to validate the presented approach. The NASA MATB-II has been used to manipulate human workload conditions and consists of four simultaneous tasks: *Tracking*, *System Monitoring*, *Resource Management*, and *Communications*. The *Tracking* task requires using a joystick to maintain a crosshair on a target, where performance

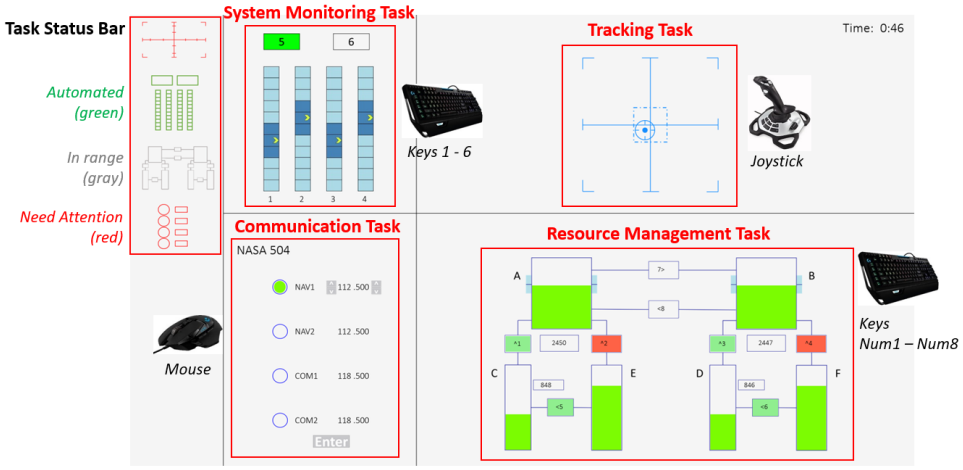

Figure 2. The NASA Multi-Attribute Task Battery-II (MATB-II) Environment.

is measured as the root mean squared error (MSE) between the target and crosshair's respective centers. Four gauges and two alarm lights were to be monitored and reset using a keyboard when out of range (gauge is too high/low and the alarm light was incorrect) in the *System Monitoring*. Performance was measured as the response time to reset an out-of-range light or gauge. The *Resource Management* task required maintaining the fuel levels of tanks A and B between 2000 to 3000 units (marked as the blue area) by using the keyboard to turn appropriate fuel pumps on/off. The percentage of time both tanks were in the range was used as the performance metric. The *Communications* task required listening and responding to audio commands. An example command is "NASA 504, NASA 504, turn you NAV1 frequency to 116.500" to which the operator responds "This is NASA 504, NAV1 set to frequency 116.500" while executing the command. Performance was measured as response time and failure rates.

Modifications were made to the NASA MATB-II environment that permitted automating each task, where automation was near perfect. Individual task status and automation status were communicated to the human operator using the status bar on the left of Figure 2. Each task icon indicated if the task needs attention (red), the task parameters are in range (grey), or the task is currently automated (green).

2) Environment Handler: The environment handler is responsible for determining what tasks to automate based on human and task state information.A reinforcement learning paradigm was used due to its ability to learn policies in a model-free fashion [20]. The task environment was defined as a Partially Observable Markov Decision Process with states S, actions A, transition $T : S \times A \times S \rightarrow [0, 1]$, set of possible observations ω, observation space $O : S \times \omega$, reward function $R : S \times A \times S \rightarrow \mathbb{R}$, and discount factor $\gamma = [0, 1]$. Each timestep was equivalent to 1 second, while selected actions (automation) were for 5 seconds.

The Soft Actor-Critic (SAC) algorithm [21] was used to decide which task to automate. The SAC algorithm interacts with the environment by taking actions to explore the state space and learn a policy to maximize the expected long-term reward and entropy. This algorithm was chosen due to its sample efficiency, as finding optimal policies online from human data is computationally expensive. The algorithm consists of an actor-network that learns the action policy and selects an action given the current state. The value (V value-

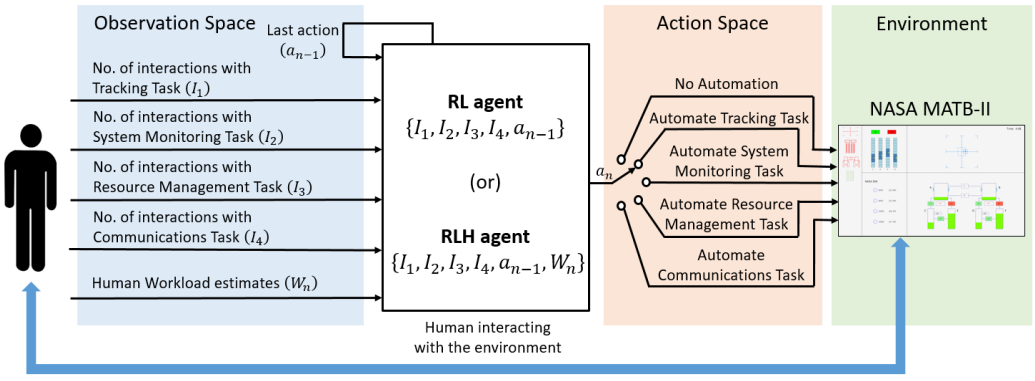

Figure 3. Soft Actor Critic agents interacting with the NASA MATB-II environment.

based) and two critic (Q value-based) networks are used to critique the action quality and suggest how to update the actor-network's policy distribution. Each network consisted of two fully connected dense layers with 64 neurons in each layer. The actor network has 2 neurons (for the mean and standard deviation of the action) in the output layer while the critic and value networks have 1 neuron each in the output layer with a linear activation function. A discount factor of 0.99 and a learning rate of 0.0003 with an Adam optimizer were used.

Two state spaces were explored with the SAC agent (see Figure 3): (i) **RL** comprising of task and interaction related information, and (ii) **RLH** augments the **RL** with human workload information as shown below:

$$S_{RL}(n) = \{I_1, I_2, \cdots, I_k, \cdots, I_K, a_{n-1}\} \quad (1)$$

$$S_{RLH}(n) = \{W_n, I_1, I_2, \cdots, I_k, \cdots, I_K, a_{n-1}\} \quad (2)$$

I_k represents the interaction information for the k^{th} task (time since last human interaction in seconds), K represents the total number of tasks, and a_{n-1} represents the last agent action. W_n represents the human workload information, which consists of estimated overall workload w_o, cognitive w_c, auditory w_a, speech w_s, visual w_v, and physical w_p workload values computed by the workload assessment algorithm. The state space is focused on interaction data than task-specific data to make the approach independent of the nature of the tasks.

The continuous action space of the SAC agent is discretized by binning the action value into $K + 1$ discrete actions (including no automation). Individual reward functions r_k are designed for each task using their performance metrics where undesirable performance is penalized and negative rewards are awarded. To prevent the agent from automating all the time and potentially disengaging the human from the system, human idle time (IT_h) is also penalized [22]. The overall reward function for the SAC agent is obtained by taking the weighted sum of the reward function of each task and human idle time:

$$r = w_0 * r_0 + w_1 * r_1 + \cdots + w_K * r_K + IT_h \quad (3)$$

Equal emphasis is given to each task by computing the normalizing coefficient for the rewards of each task obtained by an expert human operator (20+ hours of experience with the environment) with no automation. The reciprocal of the magnitude of the average reward of each task (i.e., $w_k = 1/|mean(r_k)|$) were used as the weights in Equation 3. Hence, the range of the overall rewards is $r \in [-4, 0]$. As the agent interacts with the environment, it learns and adapts its action policy based on the trends between human states and team performance to maximize the expected long-term reward.

IV. Experimental Design

Nine participants (5 males and 4 females; average age of 26.3) were recruited for the institutional review board-approved study. The mixed design experiment manipulated workload (underload (UL), normal load (NL), and overload (OL)) as its within-subjects variable and the reinforcement learning agent type (RL and RLH) as the between-subjects variable. The experiment consisted of a 15-minute training session, followed by a 52.5 minute trial with a rule-based (RB) automation strategy [18] in order to permit comparing to typical adaptive automation systems. The RB system invoked automation for the least interacted task when the human was overloaded and revoked all automation when underloaded.A five-minute break then occurred to allow the participants to rest prior to completing a 52.5 minute trial with either a RL or RLH agent, assigned randomly. Each 52.5-minute trial contained seven consecutive 7.5-minute workload conditions (OL-UL-OL-NL-UL-NL-OL), such that participants experience every workload transition. Participants completed the NASA Task Load Index (NASA-TLX) after each trial and rated their comfort and trust levels on a Likert scale from 1 (little to no comfort/trust) to 7 (complete comfort/trust).

The SAC agents were pre-trained with expert human data. The expert human data was collected while going through the 52.5-minute session 6 times with the RLH agent. Each participant started with the same pre-trained networks. The agents were then trained online without exploration during the RB trial in order to capture participant-specific data. The first 29.5 minutes of the RL/RLH trial further trained the agent and permitted the agent to explore the state space. The performance was evaluated on the last 23 minutes of data during which the RL and RLH agents were not being trained to get a consistent behavior from the agents.

V. Results & Discussion

Data from 9 participants was collected to evaluate the performance of the presented human-aware system. Two participants' data were excluded due to sensor failure. The average estimated workload, rewards, and individual task performance metrics are presented in Table I by agent type (RB, RL, and RLH). The RL agent achieved the highest overall rewards, but also the highest estimated workload. The RB approach achieved lower rewards and the lowest workload. The RLH agent maintained a lower workload level but achieved lowest rewards. Overall, the RLH agent mitigated the underload and overload workload states the best, where the most significant improvement was observed in the overload condition. The RL agent achieved higher performance than the RB and RLH agent in the system monitoring and resource management tasks, while the RLH agent performed better in the tracking and communication task. RLH agent's better performance in the communication task may be due to its decision of automating communications task most frequently. RB's performance metrics lied between RL and RLH for all four tasks.

Automation and interaction times permit analysis of the agent's decisions in relation to the human's decisions. The RB approach automated the resource management task the most, as shown in Table II. The RL agent automated the system monitoring task the most with more evenly distributed automation times amongst the tasks. Conversely, the RLH agent automated the communications task the most and did not automate each task evenly. The participants reported that the communications task was the most difficult task. The higher

Table I
MEAN PERFORMANCE METRICS WITH RB, RL AND RLH ADAPTIVE AUTOMATION APPROACHES FOR THE HUMAN-ROBOT TEAMING ON NASA-MATB ENVIRONMENT. BEST VALUES AMONG EACH TRIAL AND WORKLOAD CONDITION ARE **BOLDED**.

Workload Condition	Trial	Workload	Rewards	Tracking Error (pixels)	Gauges RT (seconds)	Lights RT (seconds)	Tanks in Range (%)	Comms RT (seconds)
Underload (UL)	RB	17.79	-1.36	**18.70**	-	**1.60**	86.41	-
	RL	**18.93**	-1.40	22.54	-	2.13	84.68	-
	RLH	17.78	**-1.30**	19.95	-	2.07	**87.52**	-
Normal load (NL)	RB	**27.26**	-1.61	25.28	3.07	3.13	84.32	12.13
	RL	33.49	**-1.54**	**22.62**	**1.80**	**1.60**	**98.14**	**11.99**
	RLH	**30.80**	-1.86	23.18	4.72	2.82	84.33	13.57
Overload (OL)	RB	35.35	-2.82	**27.44**	4.17	4.10	73.98	12.84
	RL	39.20	**-2.30**	28.51	**2.82**	**2.96**	**87.00**	12.43
	RLH	**32.67**	-3.03	**27.64**	3.29	3.55	75.00	**11.48**
Overall	RB	**26.80**	-1.93	**23.81**	4.01	3.79	81.57	12.72
	RL	30.54	**-1.74**	24.58	**2.69**	**2.67**	**89.94**	12.38
	RLH	**27.08**	-2.06	**23.59**	3.51	3.30	82.28	**11.98**

overall workload induced by the communications task may be due to the need for the human to pay attention to the audio commands, change the radio and frequency using the mouse, and verbally acknowledge the commands. This suggests that the RLH agent may have picked up this trend to reduce human workload, however it could not achieve rewards due to a much more complex state space compared to the RL agent.

The RL agent's average automation time was 917.32 seconds compared to the 637.12 seconds and 1198.75 seconds for the RB and the RLH agent automation time. This indicates that very high or very low automation time does not imply better overall team performance. The interaction time for the tracking task was the highest due to its continuous nature. The participants for the RLH trial had fewer average interactions with the system across all four tasks may indicate participants relyied on the automation more.

These results shed light on a very critical issue regarding the use of reinforcement learning for human-aware systems. Despite having more information about the human workload, the RLH may have achieved worse rewards due to the addition of the multi-dimensional human workload states which makes the state space more complex to explore. This issue may potentially be addressed by training the agent for a longer period. Another possible solution is to make the state space simpler by discretizing the workload components.

Table III shows the average NASA-TLX ratings recorded after each trial of the experiment. These ratings span six subscales: mental demand, physical demand, temporal demand, performance, effort, and frustration. The participant weights each scale in a pair-wise fashion, where the weights are used to generate an overall workload rating. The overall perceived workload with the RL agent is lower than the perceived workload with the RLH agent, which is contrary to the estimated workload based on the objective measures in Table I. The results show that the perceived performance (same trend as the rewards) biased the subjective ratings for the perceived workload. The participants indicated that Mental Demand, Performance, and Temporal Demand are the most important workload contributors during RB and RL trials while Mental Demand, Temporal Demand, and Frustration were the most important workload contributors during RLH trials.

The participants rated their trust and comfortability after each workload condition ($\sim$ every 7.5 minutes) on a 7-point Likert scale. The RL agent received the highest average trust and comfort scores of 6.07 (0.79) and 5.78 (0.62), respectively. These scores were followed by the RB approach with average

Table II
AUTOMATION AND HUMAN INTERACTION TIMINGS WITH RLH AND RL ADAPTIVE AUTOMATION APPROACHES. MOST AUTOMATED TASKS DURING EACH WORKLOAD CONDITION ARE REPRESENTED BY **BOLD VALUES**.

Workload Condition	Trial	Automation Time (seconds)				Interaction Time (seconds)			
		Tracking	Sys.Mon.	Res.Man.	Comms.	Tracking	Sys.Mon.	Res.Man.	Comms.
Underload (UL)	RB	**58.28**	15.00	10.57	19.71	137.28	0.85	6.28	0.14
	RL	89.33	**110.33**	80.00	23.00	114.33	6.66	12.00	0.00
	RLH	60.75	87.50	100.50	**156.00**	147.75	0.75	4.75	0.00
Normal load (NL)	RB	**60.57**	45.57	**59.71**	50.85	136.42	7.14	6.71	9.14
	RL	**108.00**	**111.00**	77.00	14.66	115.66	12.33	6.00	9.33
	RLH	61.75	64.00	127.75	**151.75**	122.5	10.25	5.0	6.0
Overload (OL)	RB	77.57	29.85	**146.28**	63.14	96.57	37.57	1.57	39.57
	RL	94.66	**106.66**	70.66	32.00	106.33	38.0	0.33	38.33
	RLH	47.50	57.50	124.25	**160.00**	93.00	39.75	0.25	29.25
Overall	RB	196.42	90.42	**216.57**	133.71	370.28	45.57	14.57	48.85
	RL	292.00	**328.00**	227.66	69.66	336.33	57.00	18.33	47.66
	RLH	170.00	209.00	352.50	**467.75**	363.25	50.75	10.00	35.25

Table III
NASA-TLX RATINGS FOR DIFFERENT TRIALS ON NASA MATB-II.

	RB	RL	RLH
Overall WL	67.8 (8.2)	49.9 (0.5)	71.2 (7.3)
Mental Demand	67.5 (20.7)	53.3 (13.1)	78.7 (21.5)
Physical Demand	40.6 (21.4)	40.0 (20.4)	45.0 (4.1)
Temporal Demand	77.5 (10.0)	43.3 (8.4)	81.2 (21.5)
Performance	43.7 (20.4)	20.0 (10.8)	60.0 (8.9)
Effort	74.3 (13.0)	68.3 (16.4)	71.2 (5.5)
Frustration	68.7 (18.9)	48.3 (11.7)	67.5 (5.4)

scores of 5.67 (1.03) and 5.78 (0.71) for trust and comfort, respectively. RLH agent received the lowest average trust and comfort scores of 5.50 (0.62) and 5.57 (0.80), respectively. Mann-Whitney U tests indicated that no significant difference existed between the agents for trust and comfort.

Trust and comfort are directly related to familiarity with the system. Since the participants go through the RB trial followed by RL or RLH, they are more familiar with the system during the RL and RLH trial. Despite this, RLH received the worst trust and comfort ratings. Trust and comfort are impacted by factors such as perceived team performance and the agent's action explainability. The rule-based (RB) approach is easily explainable as it is defined by a set of simple rules, but SAC agent policies as not explainable. The RL agent gained the highest trust and comfort scores despite being unexplainable as its policy led to the highest rewards.

In hindsight, the discretization of continuous action value to discrete action space may have contributed to the sub-optimal performance of the RLH agent. For future work, this work can be improved by using a discrete action space (such as SAC Discrete). Other limitations of this work include the use of the stochastic policy during the evaluation period. This means that the agent may take random actions at times which reduces explainable and performance. Future work will use deterministic policies to better assess the system.

VI. CONCLUSION

This work presented a human-aware decision-making system of systems that incorporates the human teammate's internal states for its automation decisions. Different components of the architecture were detailed and validated on the NASA MATB-II multitask environment. The human workload was estimated as the human internal state for this study; however, other human states such as fatigue, stress, and comfort can also be used. The Soft Actor-Critic reinforcement learning algorithm was used to learn the trends and relationships between the workload state and the overall team performance. Developing a human-aware system architecture can help in achieving a more fluent team collaboration by fine-tuning the robot's behavior to a human teammate's internal states; thus improving the team performance in dynamic multitask environments. The presented approach can enable a robot system to leverage human states and take the field of human factors and human-robot collaboration one step further.

REFERENCES

[1] T. B. Sheridan, "Human–robot interaction: Status and challenges," *Human Factors*, vol. 58, no. 4, pp. 525–532, 2016.

[2] R. N. Roy, N. Drougard, T. Gateau, F. Dehais, and C. P. C. Chanel, "How can physiological computing benefit human-robot interaction?" *Robotics*, vol. 9, no. 4, 2020.

[3] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): Results of empirical and theoretical research," *Advances in Psychology*, pp. 139–183, 1988.

[4] J. Heard, R. Heald, C. E. Harriott, and J. A. Adams, "A diagnostic human workload assessment algorithm for supervisory and collaborative human-robot teams," *ACM Transactions on Human-Robotic Interaction*, vol. 8, no. 2, pp. 1–30, 2019.

[5] M. R. Al-Mulla, F. Sepulveda, and M. Colley, "A review of non-invasive techniques to detect and predict localised muscle fatigue," *Sensors (Basel)*, vol. 11, no. 4, pp. 3545–3594, 2011.

[6] C. D. Wickens, J. D. Lee, Y. Liu, and S. E. G. Becker, *An Introduction to Human Factors Engineering*, 2nd ed. Pearson Education, Inc., 2004.

[7] J. McCraken and T. Aldrich, "Implications of operator workload and system automation goals," U.S. Army Research Institution, Tech. Rep. ASI-479-024-84B, 1984.

[8] J. R. Comstock and R. J. Arnegard, "The multi-attribute task battery for operator workload and strategic behavior research," NASA Langley Research Center, Tech. Rep. NASA Tech. Memorandum 104174, 1992.

[9] D. Mukherjee, K. Gupta, L. H. Chang, and H. Najjaran, "A survey of robot learning strategies for human-robot collaboration in industrial settings," *Robotics and Computer-Integrated Manufacturing*, vol. 73, p. 102231, 2022.

[10] S. H. Liu, C. B. Lin, Y. Chen, W. Chen, T. S. Huang, and C. Y. Hsu, "An EMG Patch for the Real-Time Monitoring of Muscle-Fatigue Conditions During Exercise," *Sensors (Basel)*, vol. 19, no. 14, Jul 2019.

[11] M. Ding, R. Ikeura, T. Mukai, H. Nagashima, S. Hirano, K. Matsuo, M. Sun, C. Jiang, and S. Hosoe, "Comfort estimation during lift-up using nursing-care robot — riba," in *2012 First International Conference on Innovative Engineering Systems*, 2012, pp. 225–230.

[12] C. Rodrigues, W. R. Fröhlich, A. G. Jabroski, S. J. Rigo, A. Rodrigues, and E. K. de Castro, "Evaluating a new approach to data fusion in wearable physiological sensors for stress monitoring," in *Intelligent Systems, BRACIS 2020*. Springer-Verlag, 2020, p. 544–557.

[13] S. Archer, M. Gosakan, P. Shorter, and J. Lockett, "New capabilities of the Armys maintenance manpower modeling tool," *Journal of the International Test and Evaluation Association*, vol. 26, no. 1, pp. 19 – 26, 2005.

[14] K. T. Durkee, S. M. Pappada, A. E. Ortiz, J. J. Feeney, and S. M. Galster, "System decision framework for augmenting human performance using real-time workload classifiers," in *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision*, 2015, pp. 8–13.

[15] S. Fuchs and J. Schwarz, "Towards a dynamic selection and configuration of adaptation strategies in augmented cognition," in *International Conference on Augmented Cognition*. Springer, 2017, pp. 101–115.

[16] A. H. Memar and E. T. Esfahani, "Objective assessment of human workload in physical human-robot cooperation using brain monitoring," *J. Hum.-Robot Interact.*, vol. 9, no. 2, dec 2019.

[17] J. Heard, C. E. Harriott, and J. A. Adams, "A survey of workload assessment algorithms," *IEEE Transactions on Human-Machine Systems*, vol. 48, no. 5, pp. 434–451, 2018.

[18] J. Heard, J. Fortune, and J. A. Adams, "SAHRTA: A supervisory-based adaptive human-robot teaming architecture," in *IEEE Conference on Cognitive and Computational Aspects of Situation Management*, 2020.

[19] C. Kothe, D. Medine, C. Boulay, M. Grivich, and T. Stenner, *Lab Stream Layer*, ver. 1.13, 2019.

[20] S. Singh and J. Heard, "Human-aware reinforcement learning for adaptive human robot teaming," in *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '22. IEEE Press, 2022, p. 1049–1052.

[21] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in *International conference on machine learning*. PMLR, 2018, pp. 1861–1870.

[22] G. Hoffman, "Evaluating fluency in human–robot collaboration," *IEEE Transactions on Human-Machine Systems*, vol. 49, no. 3, pp. 209–218, 2019.