Computers in Human Behavior 139 (2023) 107524

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Teaching agents to understand teamwork: Evaluating and predicting collective intelligence as a latent variable via Hidden Markov Models

Michelle Zhao [a,*,1], Fade R. Eadeh [b,*,1], Thuy-Ngoc Nguyen [c], Pranav Gupta [d], Henny Admoni [a], Cleotilde Gonzalez [c], Anita Williams Woolley [e]

[a] Robotics Institute, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA, 15213, United States of America
[b] Department of Psychology, Seattle University, 1215 E Columbia St, Seattle, WA 98122, United States of America
[c] Social and Decision Sciences, Dietrich College of Humanities and Social Sciences, 5000 Forbes Avenue, Pittsburgh, PA 15213, United States of America
[d] Gies College of Business, University of Illinois at Urbana-Champaign, 1206 S 6th St, Champaign, IL 61820, United States of America
[e] Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA 15213, United States of America

A R T I C L E I N F O

Keywords:
Collective intelligence
Machine learning
Human-autonomy teams

A B S T R A C T

Rapid growth in the reliance on teamwork in organizations, coupled with advances in artificial intelligence, has fueled increased use of Human Autonomy Teams (HATs) involving the collaboration of humans and agents to complete work. Although there are many successful examples of HATs, researchers and technology developers can see additional applications if agents were better able to understand the mental states of humans to anticipate what a team is likely to do next. Creating this capability requires the creation of models of team interaction that enable agents to interpret a team's current state and anticipate its future state. To build this model, we draw on research on collective intelligence (CI), which shows a team's capability to work together can be characterized by a latent collective intelligence factor, based on observations of work across a range of tasks, and which predicts a team's ability to accomplish a wide range of goals in the future. While some work uses a specific battery of CI tasks, more recent studies have identified observable collaborative process metrics that can be captured passively. Building on this work, we propose a method of evaluating CI by representing it as a latent variable represented by the hidden state in a Hidden Markov Model. The observations used as input to the model are the team's observable collaborative process behaviors (i.e., collective effort, use of task-related skills, and task-strategy efficiency). We show by learning the set of hidden states representing a team's observed collaborative process behaviors over time, we both learn information about the team's CI, predict how CI will evolve in the future, and suggest when an agent might intervene to improve team performance. Based on the model's observations, we discuss how it can help agents diagnose teamwork and possibly make interventions to improve CI by identifying areas of collaborative process (collective effort, skill use, or task strategy) that could be improved.

1. Introduction

During the past decade, rapid advances in computational science and artificial intelligence (AI) have given rise to *human autonomy teams* (HATs), or teams that involve at least one human working interdependently with at least one agent. In HATs, an agent is a computational subsystem partially or completely autonomous with respect to some aspect of collective activity, such as completing tasks, making decisions, or communicating information (Demir et al., 2016; O'Neill et al., 2020). While literature on human–computer interaction, including in HATs, has examined how humans respond to AI teammates (Glikson & Woolley, 2020; Musick et al., 2021), there is increasing recognition that humans and agents in HATs need to develop shared cognition to be successfully collaborative (Schelble et al., 2022; Wiltshire et al., 2017). Such shared cognition requires agents to understand individual human mental states and predict their future behavior. Even further, to operate as a full collaborator in an HAT, an agent needs to understand the *team's* current collaboration state, and accurately predict future states in order to anticipate opportunities to contribute or intervene to improve collaboration.

* Corresponding authors.
E-mail addresses: mzhao2@andrew.cmu.edu (M. Zhao), feadeh@seattleu.edu (F.R. Eadeh).
[1] Shared First Authorship.

https://doi.org/10.1016/j.chb.2022.107524
Received 3 June 2022; Received in revised form 16 September 2022; Accepted 9 October 2022
Available online 18 October 2022
0747-5632/

Recent work on collective intelligence (CI) provides a foundation for modeling collaborative states in HATs. Research on CI in human teams has identified a single latent statistical factor to describe the ability of a team to work together on a variety of tasks (Riedl et al., 2021; Woolley et al., 2010). To measure team CI, many existing studies use a performance-based measure calculated on the basis of teams' scores on a specific set of collective tasks. The resulting CI scores have been shown to predict future performance in a variety of teams and settings, including software programmers, student course project teams, online multiplayer video games, military crews, and management consultants (Engel et al., 2015; Kim et al., 2017; Riedl et al., 2021; Woolley & Aggarwal, 2020). More recently, researchers have identified observable metrics of collaborative processes that are significant predictors of CI, and can be treated as alternative measures which can be captured as teams work on almost any task, allowing unobtrusive and ongoing measurement of CI (Gupta et al., 2019; Riedl et al., 2021). The ability to track a real-time measure of CI opens the possibility for an agent to gauge a team's CI and potentially identify opportunities to intervene to improve it. While extant work has shown that collective intelligence remains relatively stable in teams in the absence of major interventions (Woolley et al., 2015), other work illustrates how interjecting technological tools and even autonomous teammates into a team can alter team design and the basic inputs into collective intelligence, leading to improvement (Glikson & Woolley, 2020; Glikson et al., 2019). As the capabilities of autonomous teammates develop further, for instance by picking up on cues related to team member mental states such as beliefs, goals, or emotions, they will likely become even more effective at intervening to improve team collaboration (Eadeh et al., 2022; Gupta et al., 2019). A real-time indicator of a team's collective intelligence also provides the possibility to model CI dynamically to explore how CI changes over time and how those changes signal future performance.

To explore the potential for agents to accurately detect a team's level of collective intelligence, we developed a Hidden Markov model (Rabiner, 1989; Schuster-Böckler & Bateman, 2007) to capture a team's current state of CI and to predict its future state. Just as traditional psychometric approaches use measured or observed variables to estimate latent factors (as has been done using factor analyses of team task scores to estimate CI in prior work), HMM also uses measures of observable behaviors to estimate a latent or hidden "state" or factor theorized to be the underlying driver of the observable behavior (Rabiner, 1989). However, while traditional factor analytic methods do not incorporate consideration of temporal sequences or changes in an underlying state over time, the task of HMMs is to find, given an observed sequence of behavior, a representation of the current "hidden" state and the probability of a future state in the next phase across a range of possibilities (Maruotti, 2011). HMMs are applied in many fields where the goal is to identify the influence of latent or hidden states based on observable data and to predict future states as they unfold sequentially across stages over time (Anderson et al., 2016; Kelley et al., 2008).

To develop our HMM, we first collected data from a study in which human participants played a search and rescue game over the internet with a pre-scripted agent. We use the data from these HATs to train a Hidden Markov model to capture the team's latent (or hidden) CI state based on observable behaviors in one phase of the game and to predict the team's future state in the next phase. We find our model achieves a fairly high level of accuracy in predicting the future CI state of HATs based on teams' observable collaborative process behavior. We then compare this model to other models, including a regression-based model and a multilayer perceptron model. We find that our HMM showed similar predictive ability to these comparison models. We discuss our results by considering ways in which agents might use such models to anticipate human behavior and adapt their own accordingly, including the possibility of identifying opportunities to intervene and improve the collective intelligence of HATs.

2. Related Work on Collective Intelligence

For several decades, a widely held perspective in teams concerned the articulation of the different ways that *any* team would need to coordinate to perform different types of tasks (such as creativity tasks versus decision-making tasks), with little consideration of the features of a *particular* team that would enable it to perform well on many different types of tasks (McGrath, 1984; Steiner, 1972). As a departure from and complement to that perspective, extant work on collective intelligence in teams (Woolley et al., 2010) investigated whether there are some groups that can consistently perform better than others and, if so, whether a measure of that capability would predict the future performance of the team. Across a number of different studies, researchers have observed evidence of a single latent factor that captures a team's collective intelligence, which is based on a factor analysis of scores from a team's performance on a number of different tasks and which can be used to predict the performance of the same team in the future (Engel et al., 2015; Riedl et al., 2021; Woolley et al., 2010). While the early body of work in this area focused on capturing a team's CI based on their performance on a specific battery of tasks, similar to an intelligence test for individuals, more recent work has identified observable collaborative process behaviors which can be unobtrusively captured as groups work on many different types of tasks or projects, and are strong predictors of standardized measures of collective intelligence (Gupta et al., 2019; Riedl et al., 2021).

In this work, we will extend these approaches to measuring collective intelligence to further explore ways that an algorithmic teammate might model and predict a team's CI based on observations over time. Hidden Markov models (HMMs) are a method for modeling sequential phenomena and have been widely applied in natural language processing (Dethlefs & Cuayáhuitl, 2011) and gene sequence analysis (Yoon, 2009). Other applications of HMMs are common in computational finance (Mamon & Elliott, 2007), machine translation (Wang et al., 2018), activity recognition (Trabelsi et al., 2013), and speech generation (Dethlefs & Cuayáhuitl, 2011). HMMs handle data represented as sequences of observations over time. The data are modeled as "observed" outputs which are theorized to have been generated by an unobserved, or "hidden," internal state. The task of HMMs is to find, given an observed sequence of behavior, a representation of the current "hidden" state and the probability of the group's future state in the next phase across a range of possibilities. HMMs are applied in many fields where the goal is to identify the influence of latent or hidden states based on observable data and to predict future states as they unfold sequentially across stages over time. HMMs have been applied in prior research to human problem-solving (Anderson & Fincham, 2013), design processes (McComb et al., 2017), and human information processing (Borst & Anderson, 2015). Using an HMM, Anderson (2011) models algebraic problem solving in the context of intelligent tutoring systems, and uses the model to predict what next problem-solving step participants will take. In a similar vein, we use our model to predict metrics of collaborative process in the next timestep.

When considering good candidates for observable behaviors to be used in an HMM analysis in order to model a group's current level of CI as well as predict its level in the next stage, we build on the work of Riedl et al. (2021). In their meta-analysis of more than 1300 teams, the researchers identified three collaborative process behaviors that are strong correlates of CI in teams: the level of collective effort of a team, the efficiency of a team's task strategy, and its use of the knowledge and skill of members. These behaviors have been identified in the classic work on team effectiveness (Hackman, 1987) as key process criteria to diagnose the quality of teamwork. The collective effort of a team can be captured by the overall level of activity of members and serves to signal the level of motivation and engagement of the team. The efficiency of a team's task strategy is captured by how effectively they use a resource, such as time and member attention, to complete work at a productive rate with a high level of quality. The use of knowledge and skill captures the degree to which a team is capitalizing on members' knowledge skills and abilities by allocating the right tasks to members with the highest relevant skill and maximizing the amount of time members spend working on tasks that use their unique skills. When teams collaborate in digital environments, it is often possible to design measures to capture these behaviors of the collaborative process and use them to predict performance (Gupta et al., 2019). In this study, we propose to use HMMs in the sequence of observed collaborative process metrics in teams playing an online search and rescue game to model the evolution of collective intelligence of a group, a latent factor influencing the behavior of the team. We use learned HMMs to predict the team's CI in future time steps, which could allow an algorithmic teammate to potentially intervene in a team struggling.

3. Method

Agents designed for collaborating in HATs must be able to judge how well a team is collaborating and anticipate what it might do next in order to operate as collaborative team members. Observable indicators of the quality of collaborative team processes that have been shown to predict collective intelligence include the appropriate team member *skill use*, the efficiency of *task strategy*, and the level of collective *effort* (Riedl et al., 2021). Extant work has shown that aggregate values of these processes predict future team behavior and performance in a variety of tasks and contexts, including search and rescue tasks (Eadeh et al., 2022), and interventions to improve targeted collaborative processes by automated agents can produce significant improvement in team CI (Gupta et al., 2019). We explore a method of interpreting and utilizing learned transition matrices to predict future CI states, which can help agents in HATs anticipate team behavior and perhaps decide when intervening in some way to improve team process could be helpful. After evaluating initial models trained with collaborative process metrics, we compare this model to other alternative models to see if HMM predictions improve the ability of agents to predict the trajectory of CI in a group and inform decisions about possible interventions.

3.1. Participants and design

We recruited 192 participants from Prolific.co, an online platform, to play a search and rescue game developed by Nguyen and Gonzalez (2022) called Minimap (see Fig. 1). Participants were assigned to a team with an ostensible teammate (in reality a pre-programmed agent, heretofore referred to as the "teammate"), who played on the same two-dimensional map as the actual participant. The field of view of the participants in the game was limited to a diameter of five squares around them, so they could not see much of the map except for a dot associated with their "teammate's" location.

Participants earned points for their team by rescuing victims that were worth either 10 or 20 points each. The total number of points the participant and their teammate earned during the two round of the game determined the team's bonus payment.

Beyond their own work in finding and saving victims, participants could also collaborate with their teammate in a number of ways, including identifying to their teammate the location of special victims only the teammate could save, communicating whether a specific location was "cleared" of victims, and activating gift boxes that would directly contribute to the bonus of teammates. These collaborative teamwork behaviors provided important input for our collaborative process metrics.

3.2. Measures

Collaborative process metrics. In modeling and predicting latent CI states for human agent teams, we use the set of collaborative process metrics identified in existing research predicting collective intelligence (Gupta et al., 2019; Riedl et al., 2021). Collaborative process metrics are computed at 30-second intervals throughout the first of the two 5-minute search and rescue missions. In the context of the Search and Rescue task, we had interval t denote minute i to $i+30$ seconds, and $P=\{1,2\}$ denote the set of teammates (one human, one agent), where $|P|$ is the total number of team members. We computed three collaborative process measures of effort, skill usage, and task strategy. Then, we operationalized the metrics as follows:

Effort was calculated as the total distance traveled by all HAT members.

$$\text{Effort at interval } t = E^t = \sum_{p=1}^{|P|}\left[\sum_{h=i}^{i+30}\sqrt{(x_{h-1}-x_h)^2+(y_{h-1}-y_h)^2}\right] \quad (1)$$

where a player's position in second h is represented as (x_h, y_h). We compute the L_2 distance between the player's location in the current second and in the previous second. Effort is summed across both human and agent teammates.

Skill use was calculated as the number of messages communicated by each participant to their partner as a means of directing partner attention to tasks that were complete versus incomplete, as well as those they were uniquely equipped to handle. For instance, in the Search and Rescue mission, some victims could only be triaged by the agent-based teammate, while others could only be triaged by the participants. Therefore, making each other aware of where their skills were needed was interpreted as an indicator of managing team skill use.

$$\text{Skill Usage at interval } t = U^t = \sum_{p=1}^{|P|}\left[\sum_{h=i}^{i+30}\text{Number of messages sent by player } p \text{ via chat at second } h\right] \quad (2)$$

Task Strategy, or the efficiency of the team's coordination, is defined as the rate of progress throughout the task, here calculated as the number of victims triaged at each point in time.

$$\text{Task Strategy (task progress) at interval } t = S^t = \sum_{p=1}^{|P|}\left[\sum_{h=i}^{i+30}\text{Number of victims triaged by player } p \text{ at second } h\right] \quad (3)$$

For each human-agent team, we compute the three process metrics of skill, effort, and strategy every 30 seconds, resulting in 10 measurements throughout a 5-minute Search and Rescue mission. These serve as observations for the HMM (Fig. 2).

4. Analytic approach

4.1. Hidden Markov Models

Hidden Markov Models (HMMs) are a formulation for learning probabilistic models of linear sequences. HMMs model a sequence of observed data as a series of outputs generated by a sequence of hidden, internal states. The observations are modeled as the output of a discrete stochastic process Observed events are often explained by hidden related variables. For example, in medical diagnosis, an illness that is classified as a "syndrome" is essentially a hidden state that is believed to be the underlying variable or driver of a set of symptoms. Consequently, when a patient is exhibiting a subset of related symptoms or experiences a change in symptoms, prediction of what might happen next is made easier by understanding their relationship to the hidden internal state or "syndrome". The process

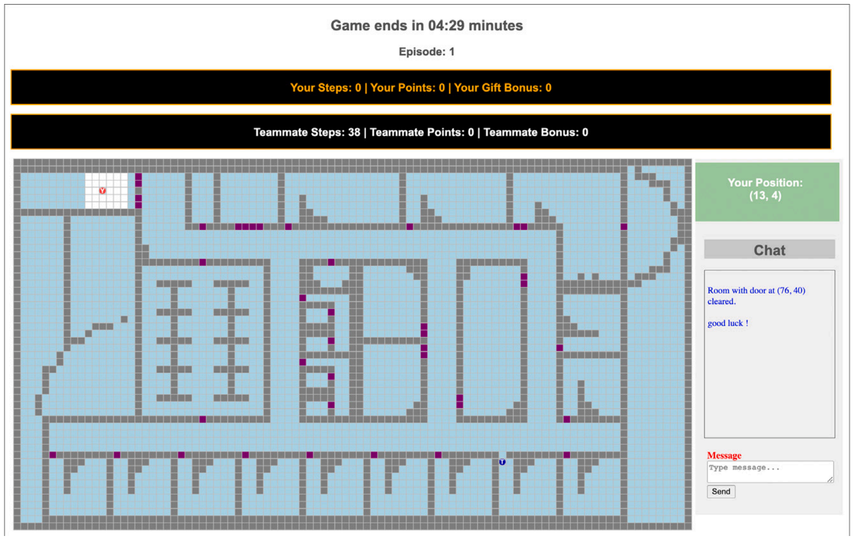

Fig. 1. Overview of the search and rescue game in Minimap. Note that participants had a field of view of only five squares at a time.

Timestep	1	2	3	4	5	6	7	8	9	10
Minute	0-0:30	0:30-1	1-1:30	1:30-2	2-2:30	2:30-3	3-3:30	3:30-4	4-4:30	4:30-5
Team 1	E^1, U^1, S^1	E^2, U^2, S^2	E^3, U^3, S^3	E^4, U^4, S^4	E^5, U^5, S^5	E^6, U^6, S^6	E^7, U^7, S^7	E^8, U^8, S^8	E^9, U^9, S^9	E^{10}, U^{10}, S^{10}

Fig. 2. Observations of process metrics. The metrics are shown for Team 1 as an example.

of identifying a medical syndrome and how it evolves over time can involve the development of models such as HMM. To develop or train the model, a variety of unstructured or semi-structured data are used as inputs (such as symptoms) from which the model derives semantic meaning by identifying connections which are conceptualized to represent underlying internal or hidden states based on patterns in how variables change over time. While the series of internal states is hidden, the model identifies the sequence of internal states (such as stages of an illness) that probabilistically generates a sequence of observations. The underlying process for hidden states is assumed to behave as a Markov process (Miller, 2001).

The *state transition* probability is the probability of moving from one hidden, internal state to another. The *emission* probabilities govern how observations (emissions) are generated from a particular hidden state. Learning the parameters of an HMM, given a set of sequences of observation, entails learning the state transition probabilities and emission probabilities. The probability of any sequence can then be computed by multiplying the state transition and emission probabilities for each entry in the sequence.

By using an HMM, we aim to model CI as a hidden temporal process that influences the sequence of observable team process metrics over time (i.e., skill use, task strategy, effort). In other words, a team's collective intelligence is a hidden state that is theorized to cause observable changes in team process metrics. Because of the theorized relationship between CI and the observable team process metrics we measure, we use measures of team process metrics as input observations to the HMM in order to encourage the HMM to represent notions of CI in the latent states. Although the hidden states may capture a number of latent attributes, we aim to gain a semantic understanding of hidden states by investigating what emissions are most probable from each hidden state. Emissions with process metric values that total a higher sum indicate that effort, task strategy, and skill use are higher on average. Given the relationship between CI and these process metrics (Gupta et al., 2019; Riedl et al., 2021), such a hidden state with above-average process metrics may be associated with greater CI than a hidden state with below-average process metrics. This allows us to loosely interpret internal states along some dimension of CI, based on probable emission values. In this way, the HMM formulation allows us to incorporate temporal information on how CI develops dynamically, and model CI as a dynamic, unobserved temporal process underlying the way in which a team's collaborative process metrics evolve over time. The HMM model provides insights that traditional auto-regressive models cannot by modeling the latent *temporal* process of CI from the *sequential* observations of team collaborative process metrics and representing how both change over time.

We model collective intelligence as a latent (hidden) state related to observed collaborative process metrics through an HMM. Our HMM is specified by the following components:

1. $X = \{x_1, x_2, .., x_N\}$ represents the state space, a set of N possible hidden states.
2. $O = \{o_1, o_2, .., o_K\}$ represents the observation space, a set of K possible observations.
3. $A = X \times X$ is a transition probability from hidden state x_i to hidden state x_j. In our HMM formulation, we consider transitions between hidden states as occurring between timesteps. $A_{ij} = \mathbb{P}(x_{t+1} = x_j | x_t = x_i)$.
4. $B = X \times O$ is a transition probability from hidden state x_i to observation o_k. $B_{ik} = \mathbb{P}(o_t = o_k | x_t = x_i)$.
5. p_0 is an initial probability distribution over hidden states.

The HMM relies on the Markov assumption, which assumes that, in order to make a prediction on a future outcome, all that is needed is information about the current state. The states before the current state have no influence on the future outcomes. The Markov assumption states that $P(x_i = c | x_1, \ldots, x_{i-1}) = P(x_i = c | x_{i-1})$.

4.2. Data processing

First, we process the study data in order to define discrete-time sequences of observations, over which to train an HMM. Training an HMM on unsupervised observed time series entails learning the transition dynamics. The transitions from hidden state to hidden state and from hidden state to observation are learned (Li & Jain, 2009). The

Timestep	1	2	3	4	5	6	7	8	9	10
Minute	0-0:30	0:30-1	1-1:30	1:30-2	2-2:30	2:30-3	3-3:30	3:30-4	4-4:30	4:30-5
Team 1	N/A	$\hat{E}^2,\hat{U}^2,\hat{S}^2,$ $\hat{E}^1,\hat{U}^1,\hat{S}^1$	$\hat{E}^3,\hat{U}^3,\hat{S}^3,$ $\hat{E}^2,\hat{U}^2,\hat{S}^2$	$\hat{E}^4,\hat{U}^4,\hat{S}^4,$ $\hat{E}^3,\hat{U}^3,\hat{S}^3$	$\hat{E}^5,\hat{U}^5,\hat{S}^5,$ $\hat{E}^4,\hat{U}^4,\hat{S}^4$	$\hat{E}^6,\hat{U}^6,\hat{S}^6,$ $\hat{E}^5,\hat{U}^5,\hat{S}^5$	$\hat{E}^7,\hat{U}^7,\hat{S}^7,$ $\hat{E}^6,\hat{U}^6,\hat{S}^6$	$\hat{E}^8,\hat{U}^8,\hat{S}^8,$ $\hat{E}^7,\hat{U}^7,\hat{S}^7$	$\hat{E}^9,\hat{U}^9,\hat{S}^9,$ $\hat{E}^8,\hat{U}^8,\hat{S}^8$	$\hat{E}^{10},\hat{U}^{10},\hat{S}^{10},$ $\hat{E}^9,\hat{U}^9,\hat{S}^9$

Fig. 3. Observations of Hidden Markov Model. The sequence of observations for Team 1 is provided as an example. The observations of the HMM at time t are o_t. $o_t = (s_t, s_{t-1})$ is a tuple of the process metrics in the current time step and in the previous time step.

observations are tuples of three collaborative process metrics: effort, skill usage, and task strategy. These process metrics are continuous variables. In order to ensure rapid online computation of the HMM, we discretize the process metrics to reduce the observational space. We compare the value of the process metric at timestep t for a given team to the mean of the metric across all teams at time t. The timestep t corresponds to data collected at the corresponding interval. The interval length was set at 30 s, in order to collect sufficient progress on effort, strategy, and skill use in each interval. The process metric value is converted to a binary value: 0 if the value is below the mean, and 1 if the value is greater than or equal to the mean. That is, the binarized effort $\hat{E}_k^t$ at time t for team k is defined as 0 if the effort is less than the mean across teams and 1 if the effort is greater than or equal to the mean across teams. Let M denote the total number of human-agent teams, the binarized effort is defined as follows:

$$\hat{E}_k^t = \begin{cases} 0 & \text{if } E_k^t < \frac{1}{M}\sum_{i=1}^{M} E_i^t \\ 1 & \text{if } E_k^t \geq \frac{1}{M}\sum_{i=1}^{M} E_i^t \end{cases} \tag{4}$$

Binarized strategy, $\hat{U}_k^t$, and skill, $\hat{S}_k^t$, are defined similarly.

$$\hat{U}_k^t = \begin{cases} 0 & \text{if } U_k^t < \frac{1}{M}\sum_{i=1}^{M} U_i^t \\ 1 & \text{if } U_k^t \geq \frac{1}{M}\sum_{i=1}^{M} U_i^t \end{cases} \tag{5}$$

$$\hat{S}_k^t = \begin{cases} 0 & \text{if } S_k^t < \frac{1}{M}\sum_{i=1}^{M} S_i^t \\ 1 & \text{if } S_k^t \geq \frac{1}{M}\sum_{i=1}^{M} S_i^t \end{cases} \tag{6}$$

4.3. State featurization

An observation at timestep t in a sequence is the condition of the team's effort, skill, and strategy at t. We lastly redefine each observation to include information from the previous timestep. The inclusion of the previous timestep in the HMM observation allows the HMM to reason about how the team's effort, skill, and strategy at timestep t may be influenced by its condition not only at time $t-1$, but also at time $t-2$. Each observation at time t, $o_t = (s_t, s_{t-1})$ is a tuple of the process metrics in the current time step and in the previous time step. An observation $o_t \in O \in \{0,1\}^6$ is represented as a six-length vector: [effort at time t, skill use at time t, task strategy at time t, effort at time $t-1$, skill use at time $t-1$, workload progress at time $t-1$] = [$\hat{E}^t$, $\hat{U}^t$, $\hat{S}^t$, $\hat{E}^{t-1}$, $\hat{U}^{t-1}$, $\hat{S}^{t-1}$]. s_t is the current-time component of observation o_t: s_t = [$\hat{E}^t$, $\hat{U}^t$, $\hat{S}^t$]. s_{t-1} is the previous-time component of observation o_t: s_{t-1} = [$\hat{E}^{t-1}$, $\hat{U}^{t-1}$, $\hat{S}^{t-1}$] (Fig. 3).

4.4. Training the HMM

Our model learns the hidden state transition matrix and emission probabilities matrix through Baum–Welch (Li & Jain, 2009). The hidden state transition probabilities represent how teams transition between different latent collaborative states. The emission probabilities inform which observed process metrics are likely given a team's latent state. We select the number of sequences of hidden states for which we achieve the best-fitting model, based on the one N which minimizes the Akaike Information Criterion (AIC) (Akaike, 1974), a measure capturing goodness-of-fit and which penalizes large numbers of parameters.

4.5. Semantic interpretation of hidden states

Training an HMM involves using a corpus of unstructured or semi-structured data from which the model identifies underlying temporal processes represented in internal hidden states. Doing so does not require any prior understanding of the semantic concepts represented by the data; however, in developing and interpreting the model, researchers can attribute meaning to the hidden states based on their probable emissions. In this case, the internal states of the HMM represent the latent CI of the team as it progresses throughout the task. In conceptualizing these internal states as collective intelligence, we are attributing semantic meaning to them by interpreting them through their probable emissions, or the observable behaviors associated with changes in the internal state. We run the Viterbi algorithm (Forney, 1973) to identify the most likely sequence of hidden states for each team's observation sequence. In a hidden state sequence, each probable hidden state underlies an observation.

Next, we group observations by the corresponding hidden state to obtain the set of observed emissions from each hidden state. Recall each observations are six values in length. We sum the values in each six-length observation vector and average the sums across all observed emissions from a given hidden state. As described above, we calculate the observed team process metrics as a single value based on the sum across the six inputs (effort, task strategy, and skill use at time t and $t-1$); therefore, the average observation value does not permit a direct interpretation regarding differences between skill use, task strategy, and effort, only whether there is a higher aggregate value across the three. This allows us to loosely understand the internal states of the model by interpreting the CI levels associated with each internal state based on the average value of observations. Although the internal states may also encode other latent attributes related to collaborative dynamics, by using the probable observations to back out the meaning of internal states, we interpret the meaning of each state along a single feature: CI. For online prediction, we use the state transition matrix to predict the level of collective intelligence at the next minute for a team.

4.6. Online prediction of process metrics

Algorithm 1 Online Process Metric Prediction

Require: Set of Team Observations J, where Team $j \in J$
Require: HMM $\mathcal{H}_J = (A_J, B_J)$
$\{o_1 = (s_1^1, s_1^0), \ldots, o_m = (s_t^m, s_t^{t-1})\} \leftarrow$ sequence of process metric observations up to time m
for $t = m, \ldots, T$ **do**
  $\{x_1, \ldots, x_t\} \leftarrow VITERBI(\mathcal{H}_J, \{o_1, \ldots, o_t\})$
  $\hat{x}_{t+1} \leftarrow \text{argmax}_{x_i}\, \mathbb{P}(x_i|x_t) = \text{argmax}_{x_i}\, A_J(x_i|x_t)$
  $\hat{s}_{t+1}^r \leftarrow \text{argmax}_{o_i \in O} \frac{\mathbb{P}(o_i|\hat{x}_{t+1})}{\sum_{o_k \in O} \mathbb{P}(o_k|\hat{x}_{t+1})} = \text{argmax}_{o_i \in O} \frac{B_J(o_i|\hat{x}_{t+1})}{\sum_{o_k \in O} B_J(o_k|\hat{x}_{t+1})}$
  $\hat{o}_{t+1} \leftarrow (\hat{s}_{t+1}^r, s_t^t)$
end for

Let H_J represent an HMM learned through Baum–Welch (Li & Jain, 2009) over a set J of team observation sequences. The learned state transition probabilities, A_J, and emission probabilities, O_J, are matrices. The agent observes a given team briefly until timestep m, gathering a sequence of process metric observations up to time m: $\{o_1 = (s_1^1, s_1^0), \ldots, o_m = (s_t^m, s_t^{t-1})\}$ (Line 1 of Algorithm 1). The agent's objective is to predict process metrics for the given team online at each subsequent timestep until the end of the mission, T. We step through the procedure for making predictions at time t. For the sequence of observations up to t, we compute the maximum likelihood sequence of hidden states $\{x_1, \ldots, x_t\}$ that generate $\{o_1, \ldots, o_t\}$ (Line 3 of Algorithm 1). Prediction of process metrics is made by first determining the most likely hidden CI state in the next timestep, $\hat{x}_{t+1}$ (Line 4 of Algorithm 1). The probability of transitioning to CI state x_i from state x_t is denoted as $\mathbb{P}(x_i|x_t)$ and is derived from the state transition probability matrix, A_J, such that $\text{argmax}_{x_i} \mathbb{P}(x_i|x_t) = \text{argmax}_{x_i} A_J(x_i|x_t)$. The maximum probability observation is selected and serves as the predicted observation, $\hat{o}_{t+1}$ (Line 5 of Algorithm 1). Recall the predicted observation is defined to be a tuple of the process metrics from the current timestep and previous timestep $\hat{o}_{t+1} = (\hat{s}_{t+1}^r, s_t^t)$. The $t+1$ component of the selected observation, $\hat{s}_{t+1}^r$, is the predicted set of process metrics for the team at the next timestep (Line 6 of Algorithm 1).

5. Results

Starting with our initial 192 trials of the human agent team, we split the data into a training set (168 trials) and a holdout test set (24 trials) to develop our model. First, we applied the HMM-based algorithm to model the progression of collective intelligence states of human agent teams throughout the search and rescue mission. We trained a single HMM with four hidden states across the dataset of sequences of team collaborative process metrics. The number of hidden states represents how many latent conditions the model maintains, or the granularity by which the algorithm can model collective intelligence. In order to determine the number of hidden states in the HMM, we perform repeated K-fold ($K = 5$) cross validation on models with increasing numbers of hidden states. For a candidate number of hidden states N, we compute the average likelihood over all team observation sequences in the training set, as well as in the test set. The AIC value for each model is $AIC = 2N - 2\ln(\hat{L})$, where N is the number of parameters (hidden states), and $\hat{L}$ is the likelihood of observations averaged over all teams in either the training or test set. We average the AIC over all 5 folds and 5 different random seeds, and select the number of hidden states which minimizes AIC; An examination of Fig. 4 (below) indicated that AIC was lowest for both the training and the holdout set at 4 states.

We will refer to this model with $N = 4$ hidden states as CI-HMM, a single HMM estimating CI of all teams. Through analysis of observations of behavior identified by extant research (Riedl et al., 2021) to be strongly related to collective intelligence – collective effort, member skill use, and team strategy – our HMM generates a learned latent variable we interpret as the collective intelligence of the team. Thus, we will interchangeably refer to hidden states as collective intelligence states (CI-states). We index the hidden states based on the mean emission value such that lower-valued states correspond to teams exhibiting below-average values of effort, task strategy, and skill use, and vice versa.

5.1. Interpretation of state and emission dynamics

The learned parameters of CI-HMM are the state transition probabilities and emission probabilities. The emission probabilities matrix represent in the $(row_i, column_j)$th entry the probability of observing observation j from hidden state i. Using our semantic interpretation method (Section 4.5), we construct an approximate ordering of the hidden states based on the average value of observations emitted. The observations are correlated with CI. The ordering aims to capture loosely some relationship between hidden states based on an estimate of CI represented in the latent attributes of each hidden state (Fig. 5).

We can also interpret the dynamics of collective intelligence of the team from one state to the next from the state transition matrix for the CI-HMM model (Fig. 6). The $(i, j)^{th}$ entries of the transition matrix in Fig. 6 refer to the probability of transitioning from state i to state j, where darker red indicates a higher likelihood that a team in the CI state indicated by the value on the x-axis in one time period will be in the CI state corresponding to the value on the y-axis in the next time period. We can see that teams exhibiting poor collective intelligence, being in State 1, are likely to remain in those states. However, as teams begin to show higher collective intelligence, such as in State 2, they have higher probabilities of transitioning to even higher CI states, exhibiting increased CI. State 3 is a highly transitional state, indicating that teams could either greatly increase (to State 4) or decrease (to State 1) their demonstrated level of CI. Teams exhibiting very high values of CI in State 4, are likely to maintain high CI in their collaboration, represented by a high likelihood of staying in state 4.

5.2. Understanding the CI dynamics of individual teams

We apply the Viterbi algorithm (Forney, 1973), a standard dynamic programming approach to calculate the maximum a posteriori probability estimate for the most likely hidden state sequence corresponding to a sequence of observations. This allows us to infer the level of collective intelligence of each team at each time step in the trial. In Fig. 7, we visualize the development of CI over time for a team drawn from the training set.

One way in which a model such as CI-HMM can be useful is for capturing what the model estimates to be a team's current CI state, but also what is likely to happen in the next time step. For instance, an autonomous teammate assisting the team represented in Fig. 7 might observe low team CI at the start of the task and look for opportunities to improve behavior. An agent could also look for instances of behavior during interval 3:00–4:00 where the team improved its CI.

5.3. Evaluating predictive accuracy of CI-HMM

One way to evaluate the accuracy of our model in estimating CI at different timesteps is to evaluate how well the (latent) learned CI state predicts the observed variables of effort, skill use, and task strategy. Algorithm 1 proposes an online inference approach for predicting at timestep t a given team's collaborative process behaviors (effort, task strategy, skill use) for the next timestep $t+1$. The algorithm predicts the most likely next-timestep $t+1$ observation, with the constraint that model's prediction of its own prior timestep matches the current observation.

Based on Algorithm 1, $\hat{o}_{t+1} \leftarrow (\hat{s}_{t+1}^r, s_t^t)$ is the predicted next observation, and $\hat{s}_{t+1}^r$ is the set of three predicted process metrics (effort, skill use, and task strategy) in the next timestep $t+1$. Let s_{t+1}^t represent the true metrics of the observed process in the next timestep. Let $s_{t+1}^t(i)$ equal the ith value in the sequence of process metrics. For example, when $i = 1$, $s_{t+1}^t(i) = s_{t+1}^t(1)$ refers to effort: the first process metric in the sequence. We compute L1-loss for predicted observations on a given team. We simulate online predictions by using data up to time t to predict the process metrics $\hat{s}_{t+1}^r$ at the next timestep $t+1$. Then, we continue using data up to time $t+1$, to predict observations for $t+2$, and so on. The loss is computed from the second timestep, as the model must receive observations in order to begin predicting next-step observations for a given team. For each team, we average the loss over each timestep in order to compute the final per-team L1-loss. $L_1 \in [0, 3]$. The maximum L1-loss is 3, since the accuracy of predicting each collaborative process metric is captured as binary variables (i.e. 1 = incorrect, 0 = correct), and thus the minimum loss is 0, if the model makes no mistakes in its prediction.

$$L_1(\hat{\mathbf{s}}, \mathbf{s}) = \frac{1}{T-1} \sum_{t=1}^{T} \left[\sum_{i=2}^{3} |\hat{s}_{t+1}^r(i) - s_{t+1}^t(i)| \right]$$

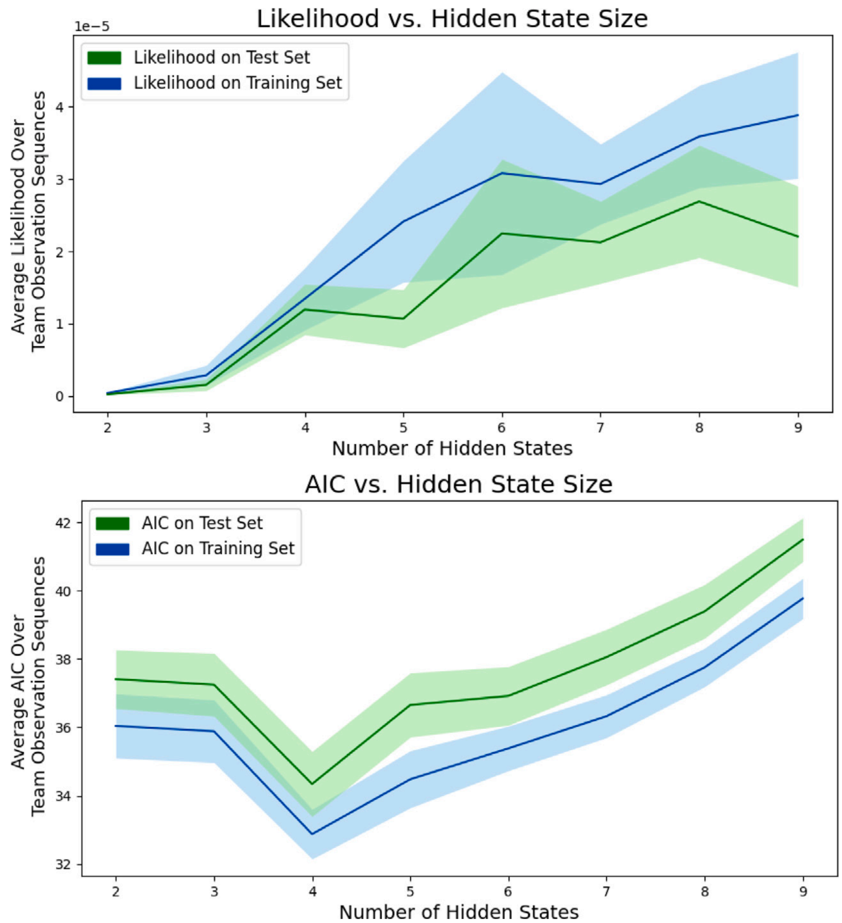

Fig. 4. We select the number of hidden states minimizing AIC. The error bars represent standard error over 5 folds and 5 different random seeds.

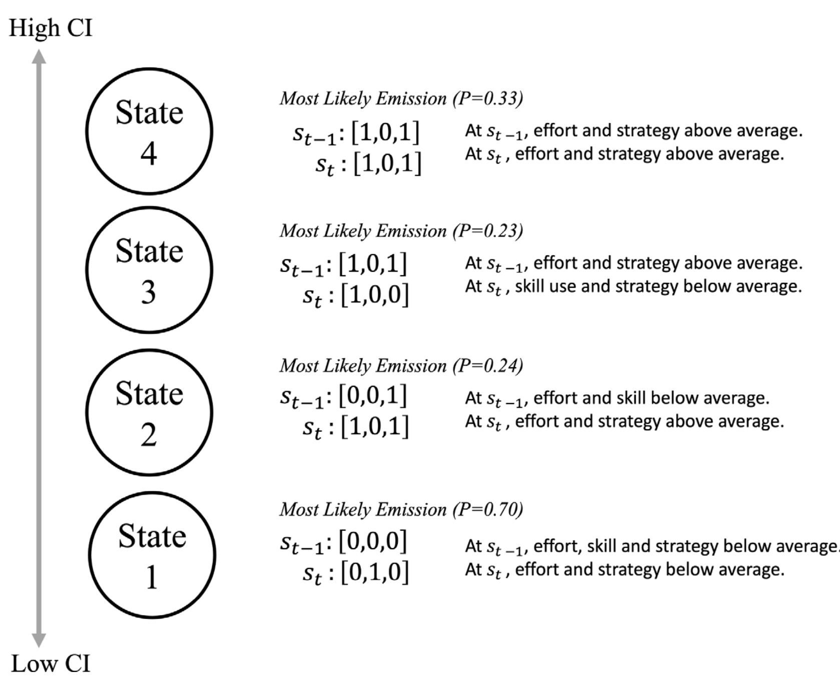

Fig. 5. Ordering of internal CI states for CI-HMM model. CI-state 4 emits observations of the highest sum on average. The most probable observation from CI-state 4 is a team demonstrating above-average effort and task strategy in previous timestep $t-1$ and continuing to do the same in timestep t. CI-state 1 is ordered as the state exhibiting the lowest CI, since it emits observations of the lowest sum on average. The most probable observation from CI-state 1 is a team demonstrating below-average effort, skill-use and task-strategy in previous timestep $t-1$ and demonstrating below-average effort and task-strategy in timestep t.

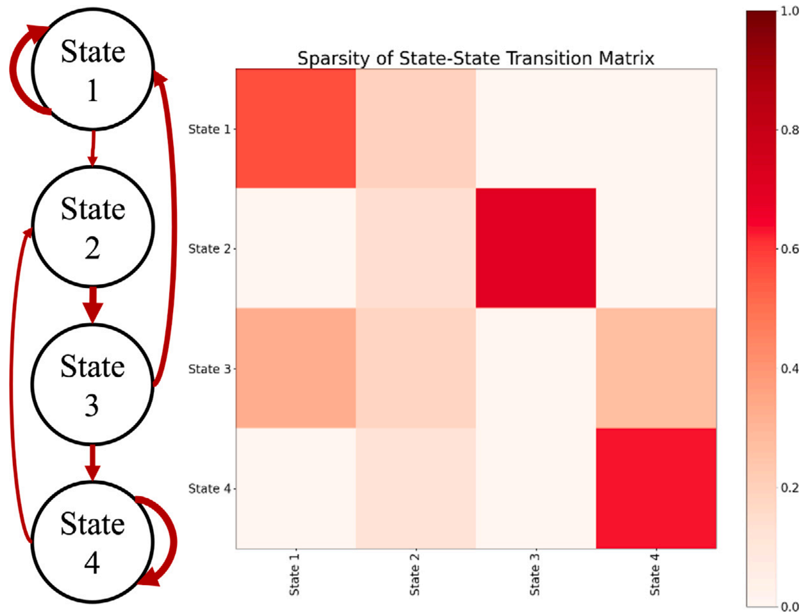

Fig. 6. State Transition Matrix for CI-HMM Model. The transition matrix provides insight into the dynamics of the latent collective intelligence variable learned by the HMM. The $(row_i, column_j)$th entries of the transition matrix refer to the probability of transitioning from state i to state j. Very poor CI states (State 1) and very high CI states (State 4) have high likelihood of transitioning back to themselves, indicating that teams with extreme CI are likely to maintain those levels. Teams in the intermediate states (States 2 and 3) have a higher spread in their probable transitions, and thus more uncertainty in how their CI will evolve. In the state diagram (left), we visualize the two most likely state transitions for each hidden state.

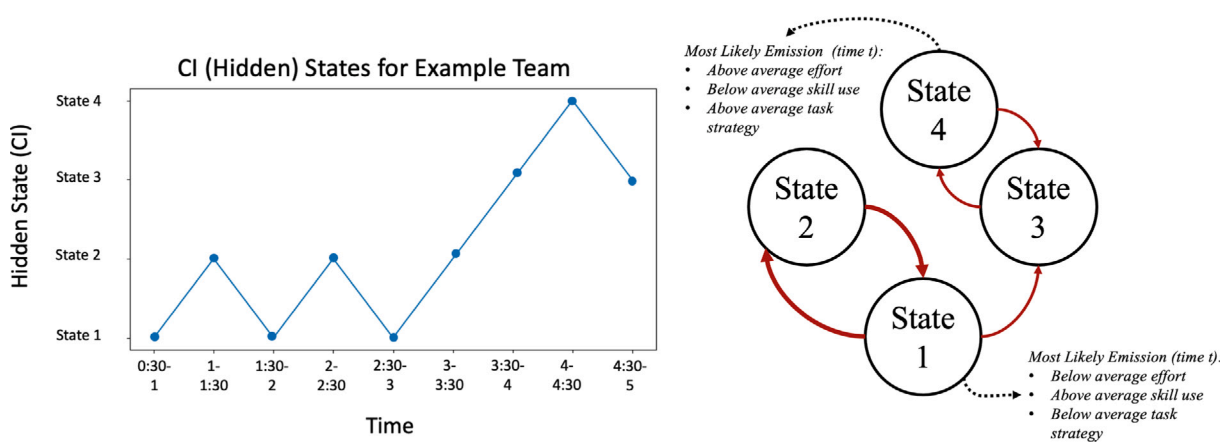

Fig. 7. The CI-HMM algorithm infers the team's collective intelligence level at each timestep. We visualize the progression of CI for a single team drawn from the training set. The line graph (left) contains the maximum a posteriori probability sequence of hidden states corresponding to the sequence of observations. Based on the CI-HMM model, we see that this team begins the mission by demonstrating mostly lower levels of CI, but transitions to higher CI levels by the end of the mission.

L1 loss for CI-HMM is calculated on the teams in the training set and separately on teams in the holdout test set (Fig. 4). Using $K = 5$ fold cross validation, we averaged the loss on training and test sets over 5 folds. We repeated the experiment over 10 random seeds, taking the model at the random seed which minimizes average training loss over the K folds. The mean training loss is 0.76 with a standard deviation of 0.05 over the 5 folds. The standard deviation is taken over the average for each random seed and fold. The mean test set loss is 0.77 with a standard deviation of 0.07.

The CI-HMM on average makes one mistake in its prediction of the team's process metric observations, predicting one of effort, skill use, or task strategy incorrectly, but the remaining two correctly, overall supporting the utility of the model in predicting future collaborative behavior. This evaluation of CI-HMM demonstrates the use of Hidden Markov Models as a candidate approach to modeling collective intelligence, a latent factor that characterizes the quality of HATs and predicts its development over time. These models are valuable for helping agents understand the dynamics of CI, and how different levels may relate to and transition to one another in order to predict future HAT behavior. Using the learned latent variable dynamics, we can make predictions about the *observed* collaborative process behaviors of human-agent teams and directly evaluate the accuracy of those predictions.

5.4. Comparison with alternative models

We also compared the HMM approach to alternative regression-based models. These include an auto-correlational linear regression (LR) model, and a multilayer perceptron (MLP) model.

The LR and MLP models take as input the vector of three process metrics at the previous timestep [effort at time t, skill use at time t, workload progress at time t] = [$\hat{E}^t$, $\hat{U}^t$, $\hat{S}^t$]. The output ground truth is

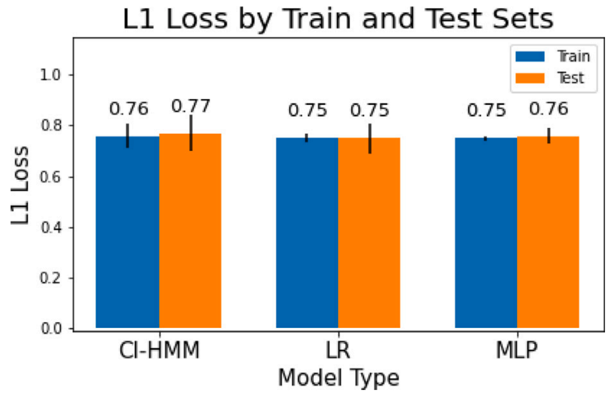

Fig. 8. The CI-HMM achieves comparable L1 loss than the linear regression (LR) and multilayer perceptron (MLP) regression models. On the training set, CI-HMM ($M = 0.76, SD = 0.05$) achieves slightly higher but similar average L_1 training loss to the LR model ($M = 0.75, SD = 0.02$) and the MLP model ($M = 0.75, SD = 0.01$). On the test set, CI-HMM ($M = 0.77, SD = 0.07$) also achieves similar average L_1 loss to the LR model ($M = 0.75, SD = 0.06$) and the MLP model ($M = 0.75, SD = 0.02$).

the vector of three process metrics at the next timestep [effort at time $t+1$, skill use at time $t+1$, task strategy at time $t+1$] = [$\hat{E}^{t+1}$, $\hat{U}^{t+1}$, $\hat{S}^{t+1}$]. We eliminate the information from time $t-1$ from the process metric vector in order to reduce correlations between elements in the input to the regression models. Since the values of the process metrics take on values in $\{0, 1\}$, we classify predicted observation values that are less than 0.5 as 0 and values that are greater than or equal to 0.5 as 1. This allows us to make a fair comparison of L_1 loss values for the vector of process metrics at time $t+1$ predicted by each model given information from time t (see Fig. 8).

For each model, we ran the model using 5-fold cross validation and compared the training and testing loss averaged over the 5-folds. For the stochastic models, CI-HMM and MLP, we tested 10 random seeds and took the minimum training and testing losses for each model. The CI-HMM model performs comparably to the linear regression and multilayer perceptron model. CI-HMM ($M = 0.76, SD = 0.05$) achieves slightly higher but similar average L_1 training loss to the LR model ($M = 0.75, SD = 0.02$) and the MLP model ($M = 0.75, SD = 0.01$). On the test set, CI-HMM ($M = 0.77, SD = 0.07$) also achieves similar average L_1 loss to the LR model ($M = 0.75, SD = 0.06$) and the MLP model ($M = 0.75, SD = 0.02$). Although the predictive performance of the CI-HMM does not outperform standard regression approaches, the CI-HMM additionally provides a representation of an underlying latent temporal process for the observed team process metrics, which standard regression approaches do not. We can use the latent stochastic process to better understand what collaborative dynamics might be underlying the data we observe.

5.5. Model extension

In building on the initial demonstration of the concept of modeling a team's hidden CI state in order to predict and/or intervene to improve future performance, there are a number of exciting possibilities for extending the model by incorporating additional inputs. One additional input to consider is information on team members' emotional or affective state, which could enhance the ability of models such as the HMM developed here to predict future CI states. For instance, two emotions that are regularly observed in settings where individuals are asked to perform challenging tasks, particularly when other humans and/or technology are involved, are anxiety and anger (Walter et al., 2014) which both have implications for team collaboration (Eadeh et al., 2022). Therefore, enabling agents to observe when human teammates are experiencing these affective states is likely to improve their ability to predict future behavior.

In exploring this possibility, we elaborated our HMM by incorporating data based on multi-item state-based measures of anger, anxiety, and positive emotion which participants completed immediately prior to the first round of the search and rescue game. We then evaluated the ability of our model to predict future states. Overall, we found incorporating data on affective state improved our model's overall accuracy, and particularly incorporating information on human teammates' levels of anger was most helpful in predicting future states, moreso than measures of anxiety and positive emotion.

We offer these observations as a means of illustrating future possibilities for elaborating similar models by incorporating additional input, not only about team member affect, but also including other inputs such as interactional synchrony (Woolley et al., 2022), linguistic analyses (Riedl & Woolley, 2017), vocal patterns (Tomprou et al., 2019), facial expressions (Chikersal et al., 2017), and other observable inputs that could be informative for predicting future team behavior.

6. Discussion

As HATs become more common in a variety of settings, there is growing recognition that machines need to be able to interpret human behavior and predict what teammates are likely to do in order to anticipate what they need and potentially make helpful interventions. In this work, we present a novel application of HMMs as a method for agents to learn a team's level of collective intelligence based on observations of sequences of team collaborative process behavior. We explore a method of interpreting and utilizing learned transition matrices to predict the future behavior of the team collaboration process. Our method enables agents to engage in real-time HAT monitoring and potentially diagnose and intervene to increase collective intelligence.

6.1. Implications for research and practice

This study represents an initial demonstration of a process for building a model to endow AI agents with a form of artificial social intelligence that can enable them to collaborate more effectively with humans. While we did not incorporate any agent-based adaptation or intervention in this study, we did observe diagnostic patterns that would provide insight to an agent about team dynamics and what they might expect regarding a team's level of effort, skill use, and task strategy, with implications for interventions they might make to help. For example, in examining the state transition matrix, we observed that teams in States 2 or 3 (equivalent to a moderate level of CI) were likely to transition to a higher level. By contrast, teams in States 1, as well as those in State 4, are much more likely to remain in their respective states of low or high CI. Consequently, an agent observing a team in States 2 or 3 could actively try to stimulate higher levels of collaborative process, particularly in any areas that seem to be noticeably weak or absent, possibly by engaging in them itself. For instance, in a team with a low level of collective effort, an autonomous teammate could engage disengaged members or could provide visualizations that demonstrate the team's level of collective effort relative to an average team. Alternatively, more effective skill use might be stimulated through noting strengths or resources of different members and encouraging the team to more explicitly consider how to put them to best use (Gupta et al., 2019).

Our findings have some important implications for research. First, we demonstrate a method for advancing the development of artificial social intelligence in the context of teamwork. Additional studies and datasets could be used to further identify observable team process behaviors that would further enhance the accuracy of the models. Furthermore, this approach opens up a whole new set of research questions pointing to exciting possibilities. How can agents use these models to successfully intervene and improve team collective intelligence? Extant work has begun to explore potential approaches to using agent-based teammates to improve team collaboration (Gupta et al., 2019; Zhou et al., 2018), demonstrating that fewer interventions that are less directive and more in the form of "nudges" seem to be most effective. Additional research can build on extant work on the best form of intervention (Maynard et al., 2020; Mohammed & Schillinger, 2022; Shuffler et al., 2018) as well as timing for team intervention (Fisher, 2017; Hackman & Wageman, 2005; Woolley et al., 2008), in addition to considering the form the input or feedback would ideally take (Eddy et al., 2013; Glikson et al., 2019; Rowe et al., 2020).

Our findings also have important implications for practice, particularly as it relates to developing AI and integrating AI agents into human collaboration. While a growing body of work in human–computer interaction and related fields has considered how AI systems can be designed to anticipate human needs and to be more explainable and transparent in order for humans to be able to understand or anticipate what they are doing (Commission, 2020; Holzinger, 2018), somewhat less attention has been given to how agents can be developed in a manner that enables them to interpret human cues and adapt their behavior. Such an ability would require further development of cognitive models and a machine theory of mind (Nguyen & Gonzalez, 2021; Rabinowitz et al., 2018), in which machines are capable of picking up on subtle and/or nonverbal cues to make inferences about what others are feeling or thinking (Baron-Cohen et al., 2001). Agents would also need to incorporate models, such as the one we have presented here, that provide a basis for understanding and predicting collaborative behavior and teamwork. Newer research is just beginning to model the way that collective cognition is formed and supports the emergence of collective intelligence (Gupta, 2022; Woolley et al., 2022) as well as the ways that AI agents might contribute to or support CI (Gupta & Woolley, 2021). Work in this domain should continue to elaborate these models to enable their incorporation into AI systems for human collaboration.

6.2. Limitations

One limitation of the development of the CI-HMM algorithm presented here relates to trade-offs we made to simplify the observation space to enable an agent to perform learning and prediction online. The HMM observations are derived from measurements of team effort, skill use, and strategy. In this work, we wanted to leverage standard HMM algorithms on a discrete observation space in order to make it tractable for an agent using the CI-HMM algorithm to perform learning and prediction online, as the agent observes the interactions and behavior of team members. In order to construct an HMM with discrete observations, we needed to heuristically discretize the observed collaborative process metrics for teams in the dataset. In doing so, we relied on the mean values across teams in the dataset for identifying whether a given team was high or low on each metric. However, this approach assumes that the values in our dataset are representative; if it turns out that our dataset is based on an unrepresentative sample where all teams collected gave extremely poor effort on the task, our model would lead an agent to erroneously conclude that a team with higher effort relative to the mean was performing well when in fact it was demonstrating low effort relative to a more globally representative sample. Alternative methods such as discretization using norms based on larger datasets would reduce dependence on comparisons to the average of any particular dataset. As work in the area progresses and more data is accumulated there will be more opportunities to improve models by using more reliable estimates of population values.

Another limitation in the current study which can be improved upon in future research is our use of pre-scripted agents as teammates collaborating with our human participants. At this stage of our work, we did not attempt to develop an agent that would adapt its behavior in response to predicted team collective intelligence levels. However, we see several opportunities to leverage our models to inform agent behavior to make them more adaptive and helpful teammates in the future. In future research, we recommend that experimenters enable agents to model and predict the collaborative behavior of the team to inform interventions aimed at increasing CI. Such interventions need not be major; indeed, targeted "nudges" in which agents make minor suggestions or use subtle primes to focus on a specific process for a team for a short time can have significant effects. Existing work (Gupta et al., 2019) suggests agent-based interventions that are too heavy-handed can easily backfire, and thus we recommend that future studies begin with small, targeted interventions that attempt to nudge teams on a specific collaborative process dimension (Gupta & Woolley, 2021; Riedl et al., 2021).

In addition, as with all research, it is important that we seek to replicate these models and our findings in additional studies and accumulate more data over time to enable more robust model development and evaluation. We also recommend that researchers consider to explore additional inputs, such as the affective variables briefly explored in our model extension, to find other variables that would further improve accuracy. Additional relevant variables could include characteristics such as cognitive diversity in the team or the level of social intelligence of the members, both of which have been shown to influence the development of CI in extant research (Aggarwal et al., 2019; Riedl et al., 2021). Incorporating more information about team member attributes in addition to team collaboration behavior into the development of group splits for training models in the future may yield additional gains in accuracy.

7. Conclusion

We propose a method for using Hidden Markov models to represent the hidden or internal state of collective intelligence in a team and to predict future collaboration and performance. We model CI as a latent variable in a Hidden Markov model using observations of the team's effort, use of members' task-related skills, and the quality of their task strategy. We illustrate that by learning the set of hidden CI states in the team we can predict how the collective intelligence will evolve in the future, allowing agents to be better collaborators in HATs and possibly even make interventions to improve team performance. We hope that this initial demonstration and the ideas we offered for future extensions will encourage more research on ways to enable more effective collaboration in HATs.

CRediT authorship contribution statement

Michelle Zhao: Conceptualization, Methodology, Software, Writing, Data curation, Visualization, Formal analysis. **Fade R. Eadeh:** Conceptualization, Methodology, Writing, Data curation, Project administration. **Thuy-Ngoc Nguyen:** Software, Data curation, Writing. **Pranav Gupta:** Writing. **Henny Admoni:** Funding acquisition, Writing. **Cleotilde Gonzalez:** Funding acquisition, Software design. **Anita Williams Woolley:** Funding acquisition, Writing, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This research is based upon work supported by the Defense Advanced Research Projects Agency, award number: FP00002636. Any opinions, findings and conclusions or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect the views of DARPA.

References

Aggarwal, I., Woolley, A. W., Chabris, C. F., & Malone, T. W. (2019). The impact of cognitive style diversity on implicit learning in teams. *Frontiers in Psychology*, *10*, 112.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.

Anderson, J. (2011). Tracking problem solving by multivariate pattern analysis and hidden Markov model algorithms. *Neuropsychologia*, *50*, 487–498.

Anderson, J., & Fincham, J. (2013). Uncovering the sequential structure of thought. *Cognitive Science*, *38*.

Anderson, J. R., Pyke, A. A., & Fincham, J. M. (2016). Hidden stages of cognition revealed in patterns of brain activation. *Psychological Science*, Article 0956797616654912, 00000.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" test revised version: A study with normal adults, and adults with asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(2), 241–251.

Borst, J., & Anderson, J. (2015). The discovery of processing stages: Analyzing EEG data with hidden semi-Markov models. *NeuroImage*, *108*, 60–73.

Chikersal, P., Tomprou, M., Kim, Y. J., Woolley, A., & Dabbish, L. (2017). Deep structures of collaboration: Physiological correlates of collective intelligence and group satisfaction. In *Proceedings of the 20th ACM conference on computer-supported cooperative work and social computing*. 00000.

Commission, E. (2020). White paper on artificial intelligence: A European approach to excellence and trust. Com (2020) 65 Final, European Commission Brussels.

Demir, M., McNeese, N. J., & Cooke, N. J. (2016). Team communication behaviors of the human-automation teaming. In *2016 IEEE international multi-disciplinary conference on cognitive methods in situation awareness and decision support* (pp. 28–34). IEEE.

Dethlefs, N., & Cuayáhuitl, H. (2011). Hierarchical reinforcement learning and hidden Markov models for task-oriented natural language generation. In *Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies* (pp. 654–659).

Eadeh, F. R., Zhao, M., Gupta, P., Nguyen, T. N., Gonzalez, C., Admoni, H., & Williams Woolley, A. (2022). The effects of anger on individual and team performance. unpublished manuscript.

Eddy, E. R., Tannenbaum, S. I., & Mathieu, J. E. (2013). Helping teams to help themselves: Comparing two team-led debriefing methods. *Personnel Psychology*, *66*(4), 975–1008, 00002.

Engel, D., Woolley, A. W., Aggarwal, I., Chabris, C. F., Takahashi, M., Nemoto, K., Kaiser, C., Kim, Y. J., & Malone, T. W. (2015). Collective intelligence in computer-mediated collaboration emerges in different contexts and cultures. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems* (pp. 3769–3778).

Fisher, C. M. (2017). An ounce of prevention or a pound of cure? Two experiments on in-process interventions in decision-making groups. *Organizational Behavior and Human Decision Processes*, *138*, 59–73.

Forney, G. (1973). The viterbi algorithm. *Proceedings of the IEEE*, *61*(3), 268–278.

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, *14*(2), 627–660.

Glikson, E., Woolley, A. W., Gupta, P., & Kim, Y. J. (2019). Visualized automatic feedback in virtual teams. *Frontiers in Psychology*.

Gupta, P. (2022). *Transactive systems model of collective intelligence: The emergence and regulation of collective attention* (Ph.D. thesis), Memory, and Reasoning, Pittsburgh, PA: Carnegie Mellon University.

Gupta, P., Kim, Y., Glikson, E., & Woolley, A. W. (2019). Digitally nudging team processes to enhance collective intelligence. In *Proceedings of the 7th ACM collective intelligence (CI) conference*. ACM.

Gupta, P., & Woolley, A. W. (2021). Articulating the role of artificial intelligence in collective intelligence: A transactive systems framework. In *Proceedings of the human factors and ergonomics society annual meeting, vol. 65* (pp. 670–674). SAGE Publications Sage CA: Los Angeles, CA, Issue: 1.

Hackman, J. R. (1987). The design of work teams ||. In J. Lorsch (Ed.), *Handbook of organizational behavior* (pp. 315–342).

Hackman, J., & Wageman, R. (2005). A theory of team coaching. *Academy of Management Review*, *30*.

Holzinger, A. (2018). From machine learning to explainable AI. In *2018 world symposium on digital intelligence for systems and machines* (pp. 55–66). IEEE.

Kelley, R., Tavakkoli, A., King, C., Nicolescu, M., Nicolescu, M., & Bebis, G. (2008). Understanding human intentions via hidden Markov models in autonomous mobile robots. In *ACM/IEEE international conference on human-robot interaction* (pp. 367–374). New York, NY, USA: ACM, event-place: Amsterdam, The Netherlands.

Kim, Y. J., Engel, D., Woolley, A. W., Lin, J. Y.-T., McArthur, N., & Malone, T. W. (2017). What makes a strong team? Using collective intelligence to predict team performance in league of legends. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing* (pp. 2316–2329).

Li, S. Z., & Jain, A. (Eds.), (2009). Baum-welch algorithm. In *Encyclopedia of biometrics* (pp. 60–61). Boston, MA: Springer US.

Mamon, R. S., & Elliott, R. J. (2007). *Hidden Markov models in finance, vol. 4*. Springer.

Maruotti, A. (2011). Mixed hidden Markov models for longitudinal data: An overview. *International Statistical Review*, *79*(3), 427–454, Publisher: Wiley Online Library.

Maynard, M. T., Mathieu, J. E., Rapp, T. L., Gilson, L. L., & Kleiner, C. (2020). Team leader coaching intervention: An investigation of the impact on team processes and performance within a surgical context. *Journal of Applied Psychology*.

McComb, C., Cagan, J., & Kotovsky, K. (2017). Mining process heuristics from designer action data via hidden Markov models. *Journal of Mechanical Design*, *139*.

McGrath, J. E. (1984). *Groups: interaction and performance, vol. 14*. Englewood Cliffs, NJ: Prentice-Hall.

Miller, D. R. (2001). Markov processes. In S. I. Gass, & C. M. Harris (Eds.), *Encyclopedia of operations research and management science* (pp. 486–490). New York, NY: Springer US.

Mohammed, S., & Schillinger, D. (2022). Translating time-based research into team interventions: An actionable, evidence-based approach. *Journal of Clinical and Translational Science*, *6*(1), Publisher: Cambridge University Press.

Musick, G., O'Neill, T. A., Schelble, B. G., McNeese, N. J., & Henke, J. B. (2021). What happens when humans believe their teammate is an AI? An investigation into humans teaming with autonomy. *Computers in Human Behavior*, *122*, Article 106852.

Nguyen, T. N., & Gonzalez, C. (2021). Theory of mind from observation in cognitive models and humans. *Topics in Cognitive Science*, Publisher: Wiley Online Library.

Nguyen, T. N., & Gonzalez, C. (2022). Minimap: an interactive dynamic decision making game for search and rescue missions. *Available at SSRN 4123762*.

O'Neill, T., McNeese, N., Barron, A., & Schelble, B. (2020). Human–autonomy teaming: A review and analysis of the empirical literature. *Human Factors*, Article 0018720820960865.

Rabiner, L. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2), 257–286, Conference Name: Proceedings of the IEEE.

Rabinowitz, N., Perbet, F., Song, F., Zhang, C., Eslami, S. A., & Botvinick, M. (2018). Machine theory of mind. In *International conference on machine learning* (pp. 4218–4227). PMLR.

Riedl, C., Kim, Y. J., Gupta, P., Malone, T. W., & Woolley, A. W. (2021). Quantifying collective intelligence in human groups. *Proceedings of the National Academy of Sciences*, *118*(21).

Riedl, C., & Woolley, A. W. (2017). Teams vs. crowds: A field test of the relative contribution of incentives, member ability, and emergent collaboration to crowd-based problem solving performance. *Academy of Management Discoveries*, *3*(4), 382–403.

Rowe, J., Spain, R., Goldberg, B., Pokorny, R., Mott, B., & Lester, J. (2020). Toward data-driven models of team feedback in synthetic training environments with GIFT. In *Proceedings of the 8th annual generalized intelligent framework for tutoring (GIFT) users symposium* (p. 136). US Army Combat Capabilities Development Command–Soldier Center.

Schelble, B. G., Flathmann, C., McNeese, N. J., Freeman, G., & Mallick, R. (2022). Let's think together! Assessing shared mental models, performance, and trust in human-agent teams. *Proceedings of the ACM on Human-Computer Interaction*, *6*(GROUP), 1–29.

Schuster-Böckler, B., & Bateman, A. (2007). An introduction to hidden Markov models. *Current Protocols in Bioinformatics*, *18*(1), A–3A, Publisher: Wiley Online Library.

Shuffler, M. L., Diazgranados, D., Maynard, M. T., & Salas, E. (2018). Developing, sustaining, and maximizing team effectiveness: An integrative, dynamic perspective of team development interventions. *Academy of Management Annals*, *12*(2), 688–724.

Steiner, I. D. (1972). *Group process and productivity*. Academic Press.

Tomprou, M., Kim, Y. J., Chikersal, P., Woolley, A. W., & Dabbish, L. A. (2019). Visual cues disrupt prosodic synchrony and collective intelligence in distributed collaboration. In *Proceedings of collective intelligence 2019*.

Trabelsi, D., Mohammed, S., Chamroukhi, F., Oukhellou, L., & Amirat, Y. (2013). An unsupervised approach for automatic activity recognition based on hidden Markov model regression. *IEEE Transactions on Automation Science and Engineering*, *10*(3), 829–835.

Walter, S., Wendt, C., Böhnke, J., Crawcour, S., Tan, J.-W., Chan, A., Limbrecht, K., Gruss, S., & Traue, H. C. (2014). Similarities and differences of emotions in human–machine and human–human interactions: What kind of emotions are relevant for future companion systems? *Ergonomics*, *57*(3), 374–386.

Wang, W., Zhu, D., Alkhouli, T., Gan, Z., & Ney, H. (2018). Neural hidden Markov model for machine translation. In *Proceedings of the 56th annual meeting of the association for computational linguistics (volume 2: short papers)* (pp. 377–382).

Wiltshire, T. J., Warta, S. F., Barber, D., & Fiore, S. M. (2017). Enabling robotic social intelligence by engineering human social-cognitive mechanisms. *Cognitive Systems Research*, *43*, 190–207.

Woolley, A. W., & Aggarwal, I. (2020). Collective intelligence and group learning. In *The Oxford handbook of group and organizational learning*. Oxford University Press.

Woolley, A. W., Aggarwal, I., & Malone, T. W. (2015). Collective intelligence and group performance. *Current Directions in Psychological Science*, *24*(6), 420–424, 00000.

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science*, *330*(6004), 686–688.

Woolley, A. W., Chow, R. M., Mayo, A. T., Riedl, C., & Chang, J. W. (2022). Collective attention and collective intelligence: the role of hierarchy and team gender composition. *Organization Science*, 1–17.

Woolley, A. W., Gerbasi, M. E., Chabris, C. F., Kosslyn, S. M., & Hackman, J. R. (2008). Bringing in the experts: How team composition and work strategy jointly shape analytic effectiveness. *Small Group Research*, *39*(3), 352–371.

Yoon, B.-J. (2009). Hidden Markov models and their applications in biological sequence analysis. *Current Genomics*, *10*(6), 402–415.

Zhou, S., Valentine, M., & Bernstein, M. S. (2018). In search of the dream team: Temporally constrained multi-armed bandits for identifying effective team structures. In *Proceedings of the 2018 CHI conference on human factors in computing systems* (pp. 1–13).