

A Model for Multi-Agent Autonomy That Uses Opinion Dynamics and Multi-Objective Behavior Optimization

Tyler M. Paine^{1,2} and Michael R. Benjamin¹

Abstract—This paper reports a new hierarchical architecture for modeling autonomous multi-robot systems (MRSs): a non-linear dynamical opinion process is used to model high-level group choice, and multi-objective behavior optimization is used to model individual decisions. Using previously reported theoretical results, we show it is possible to design the behavior of the MRS by the selection of a relatively small set of parameters. The resulting behavior - both collective actions and individual actions - can be understood intuitively. The approach is entirely decentralized and the communication cost scales by the number of group options, not agents. We demonstrated the effectiveness of this approach using a hypothetical ‘explore-exploit-migrate’ scenario in a two hour field demonstration with eight unmanned surface vessels (USVs). The results from our preliminary field experiment show the collective behavior is robust even with time-varying network topology and agent dropouts.

I. INTRODUCTION

This paper reports a new hierarchical architecture for modeling multi-robot autonomous systems: opinion formation at the higher group-level, and multi-objective behavior optimization at the individual level. Using previously reported theoretical results, we show the performance of the multi-robot system (MRS) can be designed by the selection of a relatively small set of parameters which reduces the complexity of performance optimization. Additionally, inter-robot communication can be limited in range and have finite bandwidth which presents a challenge for the design of an autonomous MRS with a large number of robots. In this formulation, the group selects choices in a completely decentralized manner that is essential for scalability to larger group sizes.

This framework for autonomous decision-making includes two major levels: group choice that occurs among networked robots, and individual decisions that are made locally. In this work we model and parameterize group choice as a non-linear dynamical opinion process [1]. Recent research has shown that this model of opinion dynamics is able to capture a wide range of functionality required of large networked systems including the ability of a group to achieve consensus and dissensus, break deadlock, and cascade important opinions across a group [2]. At the individual level, robots select values of heading and speed - or any other decision variable - that maximizes utility in the sense of multi-objective behavior optimization [3]. This optimization process generates a desired trajectory, or a reference input.

¹Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA tpaine@mit.edu, mikerb@mit.edu

²Woods Hole Oceanographic Institution Woods Hole, MA 02543, USA

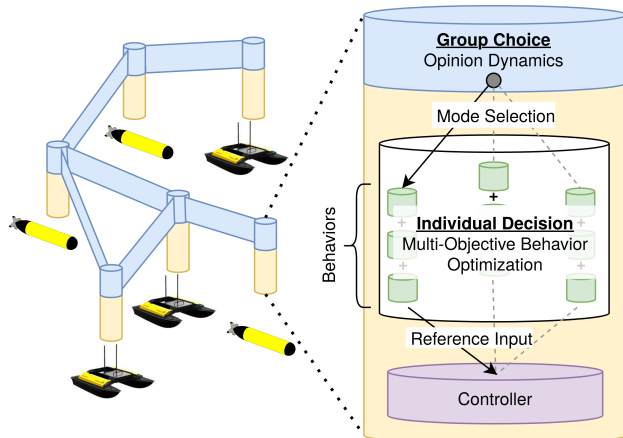


Fig. 1. Overall framework of group choice with individual decision (GCID). A networked system of robots is shown on the left, each sharing opinions with neighbors (blue). The group choice selects which local behaviors are active and the behavior optimization decides the reference input to the lower-level controller (yellow).

Although not discussed in detail in this paper, a controller is used to drive the actual state of the vehicle along the desired trajectory.

Contributions described in this paper include:

- 1) A new architecture, group choice with individual decision (GCID), for explainable decentralized multi-agent autonomy that combines a group-level choice modeled as a dynamical system of opinions with an individual-level decision making via interval programming (IvP)
- 2) A two hour field demonstration of this approach using a fleet of eight unmanned surface vessels (USVs) in an explore-exploit-migrate scenario, along with Monte-Carlo simulation studies.

II. RELATED WORK

It is widely accepted that networked multi-robot systems are more capable of completing tasks more efficiently as well as being more robust to failures [4], [5]. However, one essential prerequisite is that the autonomous population must exhibit collective intelligence [6]. A group has collective intelligence if it is able to complete problems such as task allocation [7] [8], adaptive sampling [9] [10], and formation assembly [11], [12] in response to environmental stimuli. Collective intelligence in multi-robot systems is typically designed using algorithms that exploit shared information, but can also emerge from the aggregation of simple actions from individuals with only a limited understanding of their immediate surroundings [13]. In general, more sophisticated group behaviors are associated with greater communication

between individuals in combination with increased sensory feedback and better reasoning or planning at the individual level [4], [5].

Collective behaviors of a population are a result of both competitive and cooperative behaviors by individuals [5], [6]. The authors recently fielded a MRS which provides a clear example of this insight where a group of USVs cooperatively found deep channels in rivers [7]. In that work, individual robots cooperated in building a map of the river using a consensus protocol, but then competed against each other for the most promising routes to explore using a market-based proposal bidding scheme. Complex collective behaviors present in biological groups of animals emerge from the opposing incentives to cooperate for safety yet compete for resources [14]. However, some MRS are limited because they are programmed to perform only competitive behaviors such as auctioning algorithms [11] or only cooperative behaviors such as adaptive sampling [10]. In this work, we aim to formulate a general model of robotic populations that is able to express a wide range of collective behaviors by including both competition and cooperation amongst individuals.

The combination of competition and cooperation behavior in MRS is necessary for coalition formation, where sub-groups of robots choose to work together amongst themselves, but not necessarily with other robots outside the sub-group. Coalition formation can be thought of as either separate from task allocation or as a process that occurs simultaneously with task allocation [4], e.g. consensus-based bundle adjustment (CBBA) [15]. There are many general approaches to solving a task allocation problem such as precisely encoding the tasks and holding an iterative auction [15], writing time and space constraints and solving a temporal logic optimization problem for individual robot policies as in [16], [17], and defining one or more reward functions and using reinforcement learning to learn policies [18]. Clearly, once a robot makes the choice to join a group there are many ways to collaborate. However, in this work we are interested in defining a model for dynamic coalition formation that can accommodate any style of task allocation and where coalitions perpetually reform throughout the life of the autonomous population.

Decentralized approaches to coalition formation are in general preferred because such systems are robust to communication drop-outs and failures of individual robots which happen periodically in the field. Decentralized Monte-Carlo tree search (MCTS) has been used for coalition formation as reported in [19], although most applications use MCTS for lower level path planning [20], [21]. However, in general the communication bandwidth required for MCTS approaches scales linearly by the number of agents (and the size of their state), whereas the communication cost for the opinion dynamics model used in this study scales linearly by the number of options, not agents. This property of the opinion dynamics model is increasingly important as the population grows in size. Finally, although not discussed in detail in this paper, MCTS methods theoretically require far greater computational resources than this new GCID approach.

III. NOTATION AND DEFINITIONS

We denote vectors as $\vec{v} \in \mathbb{R}^n$. The n -dimensional vector of all ones is $\vec{1}_n$. We define the strictly ordered set of mutually orthogonal standard unit vectors in n -dimensional space with $n \geq 2$ as $E(n, <)$ where

$$E(n, <) = \{\vec{e}_i : \text{where } e_i < e_j \text{ with } i < j\}$$

The state for the i^{th} agent is $\vec{x}_i \in \mathbb{R}^n$ which includes kinematic states (pose and velocities) in addition to other states such as fuel level. The state vector \vec{x} importantly does not contain information about the *opinion state* of the agent which is defined in Section IV. The connections between agents in networked systems is represented by an unweighted adjacency matrix \bar{A} with entries \bar{a}_{ik} that includes self-loops. $\bar{a}_{ik} = 1$ if the i^{th} agent is connected to the k^{th} agent and $\bar{a}_{ik} = 0$ otherwise. The set of agents connected to the i^{th} agent is Ω_i

IV. GROUP CHOICE VIA OPINION DYNAMICS

In the group choice individual decision (GCID) framework, group choice is modeled as a dynamical system of opinions. The opinion formation process is decentralized and robust to time-varying changes in network topology. Each agent holds an opinion of N_O options using the following notation adopted from the work of Anatasia *et al.* [1]: Opinions of the i^{th} agent are modeled as a vector \vec{z}_i , where z_{ij} is the opinion of the i^{th} agent about the j^{th} option, and the opinions of each agent sum to zero. More formally, $\vec{z} \in \vec{1}_{N_O}^{-1} \subset \mathbb{R}^{N_O}$. The projection onto $\vec{1}_{N_O}^{-1}$ is defined as $P_0 = I_{N_O} - \frac{1}{N_O} \vec{1}_{N_O} \vec{1}_{N_O}^T$. As defined in [1] positive values for z_{ij} imply the i^{th} vehicle has a positive opinion about the j^{th} option.

We use the Heterogeneous Inter-option Coupling model from [1] for a population of N_a agents:

$$\dot{\vec{z}}_i = P_0 F_i(\vec{z}_i), \quad (1)$$

$$F_{ij} = -d_i z_{ij} + u_i \sum_{l=1}^{N_O} S \left(\sum_{k=1}^{N_a} A_{ik}^{jl} z_{kj} \right) + b_{ij}, \quad (2)$$

where the adjacency tensor is denoted as $A \in \mathbb{R}^{N_a \times N_a \times N_O \times N_O}$ with entries A_{ik}^{jl} that parameterize the influence from the k^{th} agent's opinion about option l on the i^{th} agent's opinion about option j . We have an intuitive understanding of the parameters in this model from previous theoretical work [1]:

- A_{ii}^{jj} is the intra-agent, same-option coupling. We restrict $A_{ii}^{jj} \geq 0$, and if $A_{ii}^{jj} > 0$ the agent's opinion about this option is self-reinforcing.
- A_{ii}^{jl} is the intra-agent, different-option coupling. This parameter is used to encode interplay between agent's own opinions.
- A_{ik}^{jj} is the inter-agent, same-option coupling.
- A_{ik}^{jl} is the inter-agent, different-option coupling.

In homogeneous systems where $A_{ik}^{j_1 l_1} = A_{ik}^{j_2 l_2} \quad \forall l_1, l_2 = 1, 2, \dots, N_O$ $l_1, l_2 \neq j$ and $A_{ik}^{j_1 j_1} = A_{ik}^{j_2 j_2} \quad \forall j_1, j_2 =$

1, 2, ... N_O , cooperation and competition between agents are parameterized by A_{ik}^{jj} and A_{ik}^{jl} as: [1]

$$A_{ik}^{jj} - A_{ik}^{jl} > 0 \rightarrow \text{Cooperation} \quad (3)$$

$$A_{ik}^{jj} - A_{ik}^{jl} < 0 \rightarrow \text{Competition} \quad (4)$$

The parameter $d_i > 0$ and u_i in (2) are the resistance of the opinion and the tunable attention parameter, respectively.

We find in both simulation and field experiments that tunable attention with saturation is particularly useful when designing group behavior that must transition from a mode of normal operation to one of emergency that requires a heightened sense of urgency. Attention for the i^{th} agent is modeled as a saturating function of the magnitude of all neighbor's opinions,

$$\tau_u \dot{u}_i = -u_i + S_u \left(\frac{1}{N_O} \sum_{k=1}^{N_a} \sum_{l=1}^{N_O} (\bar{a}_{ik} z_{kl})^2 \right), \quad (5)$$

where τ_u is the time constant and S_u is the Hill saturating function with parameters as described in [1].

Since this formulation is completely decentralized, the input b_{ij} for each opinion is computed using only locally known information. This knowledge includes the state of the agent and any information shared between adjacent agents. However, to prevent redundancy with the parameters entering into the dynamical systems (2) and (5) the input b_{ij} is defined to be a utility function of only the states \vec{x}

$$b_{ij} = g_j(\vec{x}_i, \vec{x}_k) \quad \forall k \in \Omega_i. \quad (6)$$

Since each agent is constantly evaluating their own inputs and the opinions of neighbors, the choice at the group-level is continuously changing throughout the lifespan of the population. To connect the group-level choice to agent-level decision-making, we define $\vec{e}_{z_{max}}$ to represent the most favorable ordered option as a unit vector by

$$\vec{e}_{z_{max}} = \min_E \left(\arg \max_{\vec{e}_i \in E} (\vec{e}_i^T \vec{z}_i) \right), \quad (7)$$

e.g for $z_2 = z_3 > z_j \quad \forall j \neq 2, 3$, then

$$\vec{e}_{z_{max}} = |0 \quad 1 \quad 0 \quad \dots \quad 0|^T. \quad (8)$$

V. INDIVIDUAL DECISION-MAKING VIA IvP

In the group choice individual decision (GCID) framework, individual decision-making is modeled as an interval programming (IvP) optimization problem [22]. The optimizer computes the pareto optimal solution for the reference trajectory given the objective functions of all active behaviors. However, in this framework, the process of group choice described in Section IV determines for each agent which behaviors are active. The remainder of this section provides a description of the IvP process within the context of the GCID framework.

Usually, a single behavior is programmed for an individual agent to achieve a singular desired outcome, although some behaviors are designed to participate in a structured task allocation problem with neighboring agents [11] [7]. A behavior maps the values of decision variables, such as

desired heading and desired speed, to a value of utility. The decision space, S_m , for each of m decision variables, r_m , is assumed to be finite and uniformly discrete, i.e. $r_m \in S_m \subset \mathbb{R}$. The q^{th} behavior generates an objective function $f_q(r_1, r_2, \dots, r_m) : (S_1 \times S_2 \times \dots \times S_m) \rightarrow \mathbb{R}$. More complicated autonomy can be expressed through a combination of several behaviors, and IvP is used to balance competing objectives.

The multi-objective optimization problem can be written as [3]

$$r_1^*, r_2^*, \dots, r_m^* = \arg \max_{r_m \in S_m \quad \forall m} \sum_{q=1}^{N_{active}} w_q f_q(r_1, r_2, \dots, r_m) \quad (9)$$

where the utility function of the q^{th} behavior is weighted by $w_q \in \mathbb{R}$ and N_{active} is the number of active behaviors. Typically, the pareto optimal decisions of $r_1^*, r_2^*, \dots, r_m^*$ are included in the reference signal $\vec{r}(t)$ sent the vehicle controller.

In this new framework, the strongest opinion as calculated in (7) selects the behaviors that are active.

$$r_1^*, r_2^*, \dots, r_m^* = \arg \max_{r_m \in S_m \quad \forall m} \vec{e}_{z_{max}}^T A_c W \begin{bmatrix} f_1(r_1, r_2, \dots, r_m) \\ f_2(r_1, r_2, \dots, r_m) \\ \vdots \\ f_q(r_1, r_2, \dots, r_m) \end{bmatrix}, \quad (10)$$

where the diagonal weighting matrix

$$W = \text{diag}([w_1 \quad w_2 \quad \dots \quad w_q]). \quad (11)$$

The matrix A_c is the mapping from group-choice options to active behaviors on individual agents. All entries $A_{c_{jq}} \in \{0, 1\}$ are freely chosen design parameters.

In this formulation, the typical rule-based selection of active behaviors - known as MODE selection in MOOS-IvP - is replaced with a dynamic process of activating behaviors that are associated with the strongest positive opinion of the agent. As discussed in Section IV, the outcome of this process depends not only upon the local utility of each option (6), but also upon the opinions of immediate neighbors.

VI. IMPLEMENTATION DETAILS

The GCID framework requires two major components: the opinion formation system and the multi-objective behavior optimization via IvP.

A. Opinion Manager Engine

We wrote the ‘‘Opinion Manager Engine’’ to manage the dynamical process of opinion formation in a completely decentralized manner that is robust to changing network topology and communication dropouts. An overview of the Opinion Manager Engine is provided in Figure 2.

On startup, the Opinion Manager Engine loads a configuration file that specified the properties of any number of possible options to be considered. All the option parameters related to coupling, activation, and input variables are exposed as configuration parameters to eliminate changes to

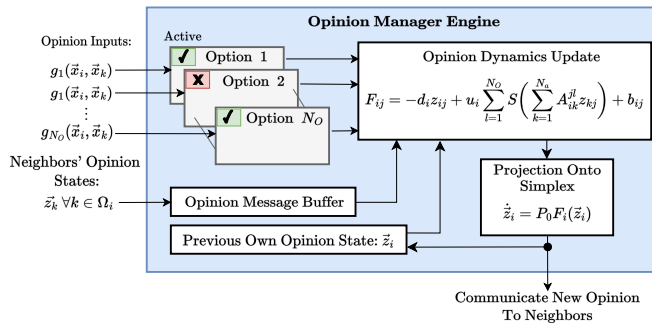


Fig. 2. Overview of Opinion Manager Engine which separately runs on each vehicle in a decentralized scheme. At each iteration, local opinion inputs, any available opinion of neighbors, and previous own opinion state are used to determine the next opinion state. The new state is communicated to any nearby vehicles.

the core code of the Opinion Manager Engine. This design choice allows the same Opinion Manager Engine to be used for any multi-vehicle mission, provided the user specifies the options with the desired parameters.

The Opinion Manager Engine handles all asynchronous communication between vehicles using a message buffer protocol where outdated opinions are ignored. Parameters related to the dynamics of option formation, such as those in (2) and (5), are specified as configuration parameters for the Opinion Manager Engine itself. The iteration frequency of Opinion Manager Engine can also be specified, and for our field testing we found a frequency of 1 Hz provided adequate response time with a relatively low bandwidth of 3-12 bytes per second depending on resolution. A MOOS App wrapper was used to connect the Opinion Manager Engine to other apps using the MOOS middleware.

B. Interval Programming (IvP) [3]

Individual behavior optimization was performed using the MOOS-IvP Helm, which is well documented in [3] and [23]. However, in this work we departed from the typical convention of hierarchical mode selection where the designer prescribes every mode via a collection of logic statements. Instead, we configured the behavior file so that some modes are selected by a rigid set of conditional statements, for example modes related to safety such as RETURN or ALL_STOP, while other modes could be chosen by the dominant opinion of the individual agent - the output of the Opinion Manager Engine shown in Figure 2. In the field experiments with USVs, the decision variables were desired speed and desired heading.

VII. EXPERIMENTAL SETUP

We demonstrated the effectiveness of this approach using a hypothetical ‘explore-exploit-migrate’ scenario that is cast as the problem of detecting and sampling algae blooms using a persistently deployed autonomous population of robotic vehicles. At the core of many scenarios in multi-agent autonomy is a tension between allocating resources to explore or to exploit. In this work we added a third option, known as migration, where a group comes to a consensus about moving to a new region and completes the move



Fig. 3. 8 Heron USVs used in a two-hour field deployment at night on the Charles River in Boston, MA

together. In this specific bloom detection scenario, the three options are described as follows:

- **Explore:** Vehicles search the region to find indication that an algae bloom was present. In this scenario vehicles are equipped with a sensor to measure turbidity in the water - since high levels are a good predictor of an algae bloom [24] - but the vehicles must move slowly for sensor readings to be accurate.
- **Exploit:** Vehicles quickly move to take samples at locations where algae blooms are indicated by vehicles that are exploring. Vehicles must stop to take a sample, which hypothetically would be transported to a lab for confirmation.
- **Migrate:** Vehicles quickly move together to a new region, usually after detecting a source of danger - such as a storm - or lack of algae blooms to sample.

An overview of the scenario is shown in Figure 4. Simulated harmful algae blooms were generated using an environmental simulator where the bloom area starts from a randomly generated location within the zone and grows outward for 10 minutes. The primary goal in this scenario was for the MRS to find the blooms and sample nearby water for confirmation. A simulated storm was randomly generated to pass over one zone, and a secondary design goal of the MRS was to avoid the storm as much as possible by moving to the other zone when a storm is detected.

A. Heron USVs

In the field experiments we used a fleet of eight Heron USVs made by Clearpath Robotics as shown in Figure 3. These USVs are designed using the frontseat-backseat paradigm, and the autonomy stack was loaded on the backseat computer. Each vehicle is equipped with a GPS and IMU sensor used for state estimation, and have wireless communication with a router at the MIT sailing pavilion. The vehicle names are: Abe, Ben, Deb, Eve, Fin, Max, Ned, and Oak.

B. Autonomy Configuration

We enumerated the options as 1 : Explore, 2 : Exploit, 3 : Migrate. The relationship between key parameters in the heterogeneous inter-option coupling model

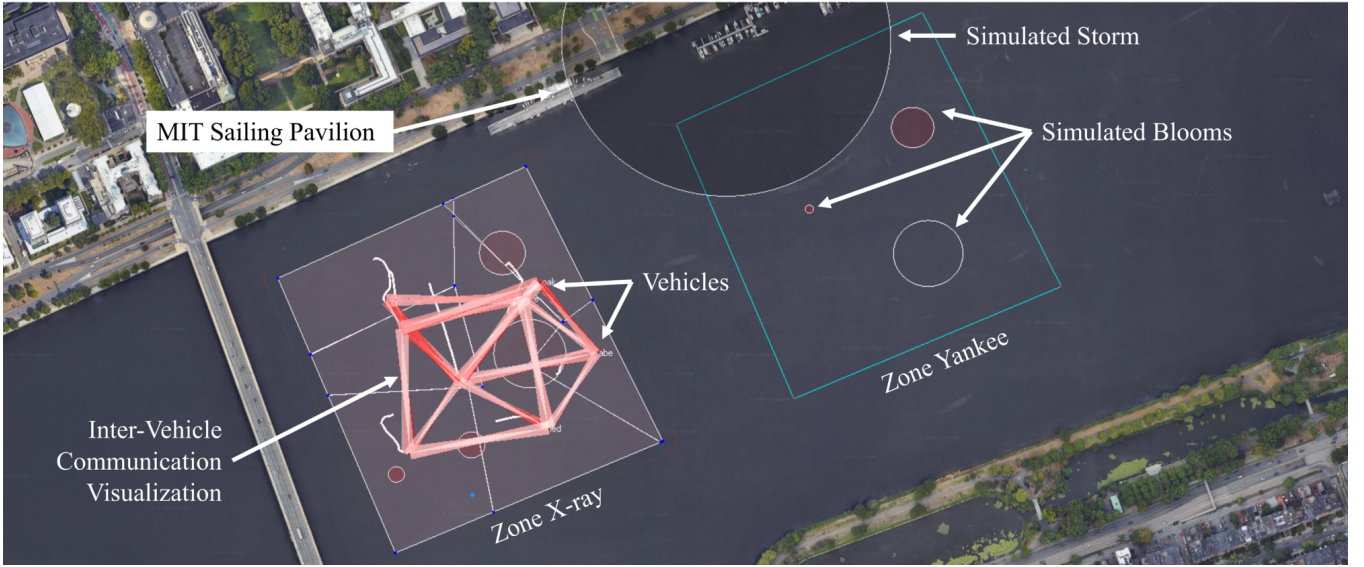


Fig. 4. State of the MRS and simulated environment approximately 50 minutes into a 2 hour mission on the Charles River near the MIT Sailing Pavilion. Simulated blooms appeared randomly within both zones X-ray and Yankee, and grew larger with time. A randomly generated storm periodically passed over the regions. At this time during the mission, 8 vehicles were searching and sampling in Zone X-ray. The communication range was artificially limited to 160 meters, and the restricted inter-vehicle communication is visualized in red. Both zones measured 300 meters by 350 meters.

(2) were:

$$A_{ik}^{jj} - A_{ik}^{jl} > 0, j = 3, \forall l, \quad (12)$$

$$A_{ik}^{jj} - A_{ik}^{jl} < 0, j = 1, 2, l = 1, 2, \quad (13)$$

$$A_{ik}^{jj} - A_{ik}^{jl} > 0, j = 1, 2, l = 3. \quad (14)$$

The design intent was to have the explore-exploit options selected primarily on a basis of social competition, while the migrate option was selected cooperative via an opinion cascade. The attention system (5) was tuned so that the social influence increased when the magnitude of the opinions increased due to a storm detection. The input b_{ij} for each option was designed as follows:

- Explore: $g_1(\vec{x}_i, \vec{x}_k)$ increases linearly with distance traveled (in simulation) or with battery level consumed (in field experiments), in addition to a positive bias when no samples are available.
- Exploit: $g_2(\vec{x}_i, \vec{x}_k)$ increases linearly with range to nearest available sample with a maximum value, in addition to a positive bias when currently sampling.
- Migrate: $g_3(\vec{x}_i, \vec{x}_k)$ includes a large positive bias when a storm is detected, in addition to a positive bias when migrating and all neighbors are not yet within the destination region. A negative bias is included after a migration is complete and it increases linearly with time since the last migration until it is equal to zero.

Due to space constraints we do not list the specific values, but they are available upon request.

Finally, since at any given time there were multiple vehicles exploring the same region and multiple vehicles traveling to multiple sample locations, two different schemes for task allocation were required. Exploring vehicles used a search algorithm based on Voronoi partitioning previously published in [25]. For sampling vehicles, an optimal route for sampling vehicles can be found by solving a distributed

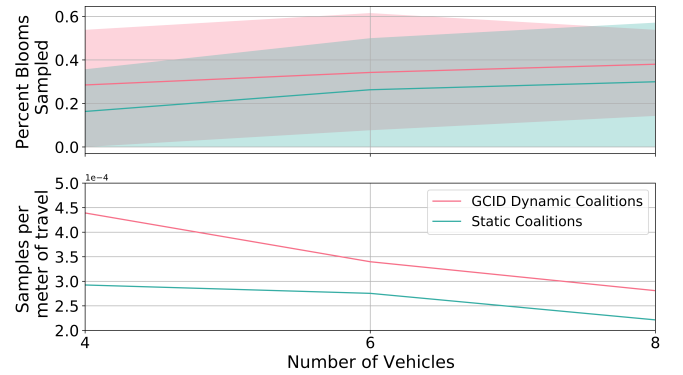


Fig. 5. Monte Carlo simulation experiments of the bloom sampling scenario with increasing fleet size. The MRS using the GCID approach sampled a higher percentage of blooms on average (top), and also sampled blooms with a slightly higher average efficiency - greater samples per meter of travel (bottom). Solid lines are the mean for each type and the shaded areas indicate the min and max values of all trials.

multiple traveling salesman problem - which is at least NP-complete [26]. In this work sampling vehicles used a greedy approach where they traveled to the closest point which was not also the closest point to another sampling vehicle.

VIII. SIMULATION RESULTS

Due to the limited amount of uninterrupted time available on the Charles River as well as the onerous supervision required for a large MRS that operated in 0.21 km^2 of busy public water-space, we conducted a series of preliminary Monte-Carlo simulation studies to evaluate the effectiveness of this method. For comparison we fixed the coalitions in the MRS such that in the absence of a storm half the fleet was exploring, and the other half exploiting. This composition of modes in the static coalition matched the collective behavior designed using the GCID framework. Simulations were completed with fleets of size 4, 6, and 8. A total of

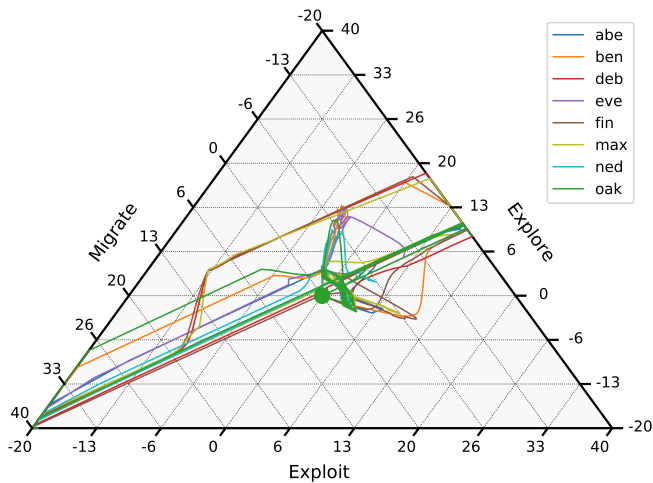


Fig. 6. Trajectories of the opinion state for each vehicle during the 2 hour field mission. Each trace shows how the opinion for a vehicle evolved over time in response to input and opinions from neighbors. This detail view on the simplex is zoomed in to the range $[-20, 40]$ to show the trajectories near the origin.

300 simulations were completed using Monte-MOOS [27], each with randomized blooms and storms.

A summary of results is shown in Figure 5. On average the GCID dynamic coalition formation outperformed the static coalitions on the basis of sampling percentage and efficiency. The same task allocation algorithms were used in both cases, and the only difference was in allowing group choice via opinion dynamics.

IX. FIELD RESULTS

We report field results from a 2 hour night-time operation on the Charles River. Using the GCID approach, all 8 USVs dynamically changed coalitions, locating possible simulated blooms and allocating vehicles to sample them. All vehicles selected to start in Zone X-ray and all migrated to Zone Yankee when one vehicle (Abe) detected a storm. A plot of the opinion trajectories is given in Figure 6 which shows two interesting features: **1)** A clear bias toward exploring during portions of the mission when there are no known locations to sample, and **2)** a bifurcation near the origin between the options to explore and exploit, with evidence that vehicle's opinion state would transition between the two options as desired.

A plot of the magnitude of the attention (5) for each vehicle is shown in the top plot of Figure 7. As designed, the periods of high attention correspond to either participating in a migration or just completing a migration. In this way, attention was tuned such that the MRS was responsive to the detection of a storm by any one vehicle, and that response was an opinion cascade for every vehicle in the network.

A. Robustness to vehicle dropout

Approximately 70 minutes into the mission, the vehicle Deb unexpectedly stopped communicating, and because a dead-man switch was implemented in software, Deb stopped the mission. The vehicle was manually driven back to the MIT sailing pavilion where it was revived and returned to

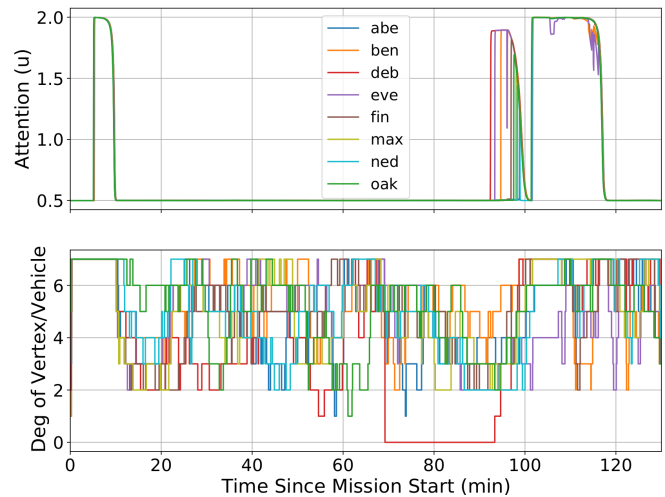


Fig. 7. Magnitude of attention (top) and degree of connection (bottom) for each vehicle throughout the 2 hour field mission. By design vehicles had high attention during or after migration events (top). The number of neighbors within communication range fluctuated throughout the mission (bottom).

the group about 25 minutes later. This event can be seen in the bottom plot of Figure 7, which shows the degree of connectivity for all vehicles - including the period when Deb maintained a degree of 0.

This unplanned event demonstrated the robustness of our decentralized approach. No additional commands were sent to the fleet while we recovered Deb, and the remaining 7 vehicles continued to dynamically form coalitions using the decentralized implementation of the Opinion Manager Engine. Furthermore, the bottom plot on Figure 7 shows that during the mission every vehicle was periodically disconnected from others. No single vehicle maintained a fully connected graph during the entire mission.

X. CONCLUSIONS AND FUTURE WORK

In this paper we reported a new combination of group choice with individual decisions (GCID) and provided preliminary evaluation of its effectiveness using both Monte-Carlo simulation trails and a field demonstration. This approach models the choice of agents in a group as a dynamical system of opinions, followed by a decision to optimize over multiple active behaviors using interval programming (IvP). Results from randomized simulation studies suggest this new method offers improved performance over static coalitions. In our field experiments we demonstrate collective behavior even with time-varying network topology and agent dropouts. Future work includes determining bounds on performance of the system as a function of the local utility functions (6), completing a more simulated comparison studies against other methods, and fielding other MRS with a different combination of vehicle types.

ACKNOWLEDGMENTS

We would like to acknowledge all members of the MIT Pavlab who participated in field testing: Mikala Molina, Mark Franklin, Karan Mahesh, Raymond Turrisi, Kevin Becker, and Spring Lin.

REFERENCES

- [1] A. Bizyaeva, A. Franci, and N. E. Leonard, "Nonlinear opinion dynamics with tunable sensitivity," *IEEE Transactions on Automatic Control*, vol. 68, no. 3, pp. 1415–1430, 2023.
- [2] A. Franci, A. Bizyaeva, S. Park, and N. E. Leonard, "Analysis and control of agreement and disagreement opinion cascades," *Swarm Intelligence*, vol. 15, no. 1-2, pp. 47–82, 2021.
- [3] M. R. Benjamin, H. Schmidt, P. M. Newman, and J. J. Leonard, "Nested autonomy for unmanned marine vehicles with moos-ivp," *Journal of Field Robotics*, vol. 27, no. 6, pp. 834–875, November/December 2010.
- [4] Y. Rizk, M. Awad, and E. W. Tunstel, "Cooperative heterogeneous multi-robot systems: A survey," *ACM Computing Surveys (CSUR)*, vol. 52, no. 2, pp. 1–31, 2019.
- [5] J. K. Verma and V. Ranga, "Multi-robot coordination analysis, taxonomy, challenges and future scope," *Journal of intelligent & robotic systems*, vol. 102, pp. 1–36, 2021.
- [6] N. E. Leonard and S. A. Levin, "Collective intelligence as a public good," *Collective Intelligence*, vol. 1, no. 1, p. 26339137221083293, 2022.
- [7] N. Gershfeld, T. M. Paine, and M. R. Benjamin, "Adaptive and collaborative bathymetric channel-finding approach for multiple autonomous marine vehicles," *IEEE Robotics and Automation Letters*, vol. 8, no. 7, pp. 4028–4035, 2023.
- [8] S. Park, Y. D. Zhong, and N. E. Leonard, "Multi-robot task allocation games in dynamically changing environments," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 8678–8684.
- [9] N. E. Leonard, D. A. Paley, F. Lekien, R. Sepulchre, D. M. Fratantoni, and R. E. Davis, "Collective motion, sensor networks, and ocean sampling," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 48–74, 2007.
- [10] S. Kemna, J. G. Rogers, C. Nieto-Granda, S. Young, and G. S. Sukhatme, "Multi-robot coordination through dynamic voronoi partitioning for informative adaptive sampling in communication-constrained environments," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 2124–2130.
- [11] M. R. Benjamin, T. Paine, and S. Randeni, "Autonomy algorithms for stable dynamic linear convoying of autonomous marine vehicles," in *OCEANS 2021: San Diego – Porto*, 2021, pp. 1–10.
- [12] W. Ren and N. Sorensen, "Distributed coordination architecture for multi-robot formation control," *Robotics and Autonomous Systems*, vol. 56, no. 4, pp. 324–333, 2008.
- [13] M. Rubenstein, A. Cornejo, and R. Nagpal, "Programmable self-assembly in a thousand-robot swarm," *Science*, vol. 345, no. 6198, pp. 795–799, 2014.
- [14] I. D. Couzin, "Collective cognition in animal groups," *Trends in cognitive sciences*, vol. 13, no. 1, pp. 36–43, 2009.
- [15] L. Brunet, H.-L. Choi, and J. How, "Consensus-based decentralized auctions for robust task allocation," *IEEE Trans. Robot.*, vol. 25, no. 4, pp. 912–926, 2009.
- [16] K. Leahy, Z. Serlin, C.-I. Vasile, A. Schoer, A. M. Jones, R. Tron, and C. Belta, "Scalable and robust algorithms for task-based coordination from high-level specifications (ScRATHeS)," *IEEE Transactions on Robotics*, vol. 38, no. 4, pp. 2516–2535, 2022.
- [17] W. Liu, K. Leahy, Z. Serlin, and C. Belta, "Catlnet: Learning communication and coordination policies from catl+ specifications," in *Learning for Dynamics and Control Conference*. PMLR, 2023, pp. 705–717.
- [18] Y.-T. Tian, M. Yang, X.-Y. Qi, and Y.-M. Yang, "Multi-robot task allocation for fire-disaster response based on reinforcement learning," in *2009 International Conference on Machine Learning and Cybernetics*, vol. 4. IEEE, 2009, pp. 2312–2317.
- [19] H. Zheng, J. Guo, X. Xie, and P. Yan, "A distributed coalition formation method of heterogeneous uav swarm in unknown dynamic environment," *Journal of Astronautics*, vol. 43, no. 2, pp. 189–197, 2022.
- [20] K. Kurzer, C. Zhou, and J. Marius Zöllner, "Decentralized cooperative planning for automated vehicles with hierarchical monte carlo tree search," in *2018 IEEE Intelligent Vehicles Symposium (IV)*, 2018, pp. 529–536.
- [21] M. Daneshvaramoli, M. S. Kiarostami, S. K. Monfared, H. Karisani, K. Dehghannayeri, D. Rahmati, and S. Gorgin, "Decentralized communication-less multi-agent task assignment with cooperative monte-carlo tree search," in *2020 6th International Conference on Control, Automation and Robotics (ICCAR)*, 2020, pp. 612–616.
- [22] M. R. Benjamin, "The interval programming model for multi-objective decision making," Computer Science and Artificial Intelligence Laboratory, MIT. Cambridge, MA, Tech. Rep. AIM-2004-021, 2004.
- [23] —, "Moos-ivp autonomy tools users manual, mit computer science and artificial intelligence lab," Computer Science and Artificial Intelligence Laboratory, MIT. Cambridge, MA, Tech. Rep. MIT-CSAIL-TR-2010-039, 2010.
- [24] M. Rome, R. E. Beighley, and T. Faber, "Sensor-based detection of algal blooms for public health advisories and long-term monitoring," *Science of The Total Environment*, vol. 767, p. 144984, 2021.
- [25] N. Evans, "A practical search with voronoi distributed autonomous marine swarms," Master's thesis, Massachusetts Institute of Technology and Woods Hole Oceanographic Institution, 2022.
- [26] S. J. Russell, *Artificial intelligence a modern approach*. Pearson Education, Inc., 2010.
- [27] K. Becker, "Development of a distributed simulation cluster for moos-ivp," Master's thesis, Massachusetts Institute of Technology, 2024, under review.