Learning Based Leadership in Swarm Navigation

Ovunc Tuzel, Gilberto Marcon dos Santos, Chloë Fleming(✉), and Julie A. Adams(✉)

Collaborative Robotics and Intelligent Systems Institute, Oregon State University, Corvallis, OR, USA
{flemichl,julie.a.adams}@oregonstate.edu

Abstract. Collective migration in biological species is often guided by distributed leaders that modulate their peers' motion behaviors. Distributed leadership is important for artificial swarms, but designing the leaders' controllers is difficult. A swarm control strategy that leverages trained leaders to influence the collective's trajectory in spatial navigation tasks was formulated. The neuro-evolutionary learning based control method was used to train a few leaders to influence motion behaviors. The leadership control strategy is applied to a rally task with varying swarm sizes and leadership percentages. Increasing the leadership representation improved task performance. Leaders moved quickly when the swarm had a higher percentage of leaders and slowly when the percentage was small.

1 Introduction

Biologically-inspired swarm robotic systems exhibit emergent behavior based on local interactions. However, coordinating swarms is challenging. Swarms' distributed, localized communication networks hinder access to global information [7], including navigation goals.

Collective behaviors in fish, birds, and bees suggests that motion coordination can be achieved by a decentralized system without global control or communication mechanisms [13]. Navigation tasks are often facilitated by distributed leaders responding to environmental stimuli [5,17–20,26,29]. The leaders are typically anonymous in large homogeneous collectives [11] and only directly influence individuals within their localized interaction neighborhoods; however, their actions propagate, creating a collective response [8,34].

Leaders play an important role in biological swarm coordination, but it is unclear how they tune their behaviors to maximize their influence over the swarm's behavior. A neuro-evolutionary learning method to train leaders is developed and evaluated in order to explore leadership mechanisms for artificial swarms. A key contribution is a learning based leadership strategy, where a simulated swarm can be influenced with leadership percentages as low as 4%.

M. Dorigo et al. (Eds.): ANTS 2018, LNCS 11172, pp. 385–394, 2018.
https://doi.org/10.1007/978-3-030-00533-7_33

2 Related Work

Collective navigation is critical for large groups and is often guided by individuals assuming leadership roles. Leadership in biological swarms can be a transient role assumed by any individual [17,21,29]. Fish trained to forage based on environmental features were inserted into a shoal of naive fish and led the shoal to the food source, even though the majority were uninformed [26]. Leadership that emerges based on internal and environmental conditions allows any swarm member to assume a leadership role, which makes them anonymous and the swarm robust to leader loss. The leader percentage in biological swarms is often small: 5% for swarming honeybees [30], and 9% for fish [26].

The mechanisms biological leaders use to guide a swarm are not well-characterized, but they can assume frontal positions [10,24]. Honeybees moving to a new hive have a few fast-flying members [5]. Frontal fish, often faster swimming food deprived individuals, have greater influence on the shoal's direction [19]. However, leaders that move too aggressively tend to leave the swarm members behind, suggesting that leaders must remain aware of their followers [16].

Learning based methods for deriving controllers are often applied to robotic swarms. Attributing global performance to individual agents' behaviors is difficult for large scale systems [33], particularly when the global state is unobservable by the individuals [4]. Several related efforts mitigate this multi-agent credit assignment problem using *team learning* [22] to train the swarm using identical controllers and reward signals for all agents [1–3,14,25].

Neuro-evolutionary team learning methods are common when generating swarm agents' controllers using a fitness function representative of collective behavior. The design paradigm emphasizes simple agent control policies based on locally observable information [7], where directly mapping sensor inputs to control actions is often suitable [2,6,23,31,32]. Several efforts [2,31,32] demonstrated that swarm aggregation tasks can be accomplished with neural network (NN) controllers using this mapping approach. Similar methods successfully generated controllers for more complex tasks (i.e., predator avoidance [28] and collaborative foraging [12]). However, influencing the swarm via distributed leaders is a complex task that has not been investigated with learning based methods.

3 Approach

A swarm of n nonholonomic homogeneous agents, $R = \{r_1, r_2, \cdots, r_n\}$, navigates a continuous 2D environment to perform a rally task, in which leaders, $L \subset R$, know the goal location and are to lead the swarm to the goal. All agents in $S = R - L$ are oblivious to the goal. Each agent controls its velocity $v \,\epsilon\, [0, v_{max}]$, where v_{max} is a constant upper limit, and desired heading angle $\psi \,\epsilon\, [-\pi, \pi]$.

All agents in S interact based on Reynolds's rules [27]: repulsion (r_{rep}), orientation (r_{ori}), and attraction (r_{att}) that delineate 2D zones around each agent, where $r_{rep} < r_{ori} < r_{att}$. These agents: (1) Veer away from all neighbors within

r_{rep}, (2) Align with neighbors at distances between r_{rep} and r_{ori}, and (3) Move towards neighbors between r_{ori} and r_{att}.

Leaders exert influence by moving among the swarm, using a policy determined by a NN to apply repulsion-orientation-attraction forces on the other agents. All leaders use the same NN, with identical weights and train using team learning.

3.1 Neuro-Evolutionary Learning Method

A set of m two-layer NNs $A = \{a_1, a_2, \cdots, a_m\}$ is initialized with random weights generated by a Gaussian distribution centered at zero, with a standard deviation of 1 ($\mu = 0, \sigma = 1$). The NNs' hidden and output layers consist of units with arc tangent activation functions to provide symmetry, and bounded the output to $[-\pi, \pi]$, consistent with the leader robots' desired heading angle, ψ.

The NN's four sensory inputs represent the polar coordinates of two points, relative to the leader's reference frame. The first pair of inputs are the polar coordinates, distance (d) and heading (θ), between the robot and the goal position, $\langle d_g, \theta_g \rangle$. The second pair of inputs are the polar coordinates from the robot to the centroid of the swarm within the leader's perceptual range, $\langle d_s, \theta_s \rangle$. All possible inputs to the NN are defined by the input vector $NN_x = \langle d_g, \theta_g, d_s, \theta_s \rangle$. The NN's outputs are the leader's desired velocity and heading $NN_o = \langle v, \psi \rangle$.

Each epoch simulates a rally task for all m NNs in the population A. An episode loads a NN a_i into the leaders L, positions the leaders and non-leaders S randomly within a starting location, and simulates a fixed number of steps τ. The NN's performance is evaluated using the cost function E (Eq. 1) upon episode completion. All m NNs are evaluated and the top-performing λ networks, called parents, are retained. A new set of $m - \lambda$ NNs are generated by randomly sampling (with replacement) from the λ parents. These $m - \lambda$ NNs are mutated by applying zero-mean Gaussian noise with a fixed standard deviation NN_{mut} ($\mu = 0, \sigma = NN_{mut}$) to every NN weight. The mutated NNs are incorporated into the evolutionary population; thus, returning the population size back to m.

At each simulation step t, the temporal factor, $\frac{t}{\tau}$ in $[0, 1]$, represents time progress over the episode's total steps τ. The temporal weight w_t is a function of the temporal factor, $w_t = 1 - cos(\pi \cdot \frac{t}{\tau})$, over its valid input range, $\frac{t}{\tau} \in [0, 1]$. The area under the curve, $w_t(\frac{t}{\tau})$, is 1.

The Euclidean distance between each non-leader agent i and the goal at step t, d_t^i, is evaluated and weighted by the temporal weight. The weighted accumulated distance is summed over every step, providing an average weighted distance $d_{avg}^i = \frac{1}{\tau} \cdot \sum_{t=0}^{\tau} d_t^i \cdot w_t$ representing the cost associated with agent i for the entire episode. The temporal weighting increases the influence of agents' deviations from the goal late in the episode. The weighting rewards NNs that consistently converge towards the goal, rather than those that initially drive the swarm towards the goal, but later disperse or lose control of the swarm.

The average accumulated weighted distance d^i_{avg} is averaged across all agents in S at the end of each episode, defining the NN fitness function:

$$E = \frac{1}{|S|} \cdot \sum_{i \in S} d^i_{avg}. \quad (1)$$

4 Experimental Design and Results

The primary research question is whether a small percentage of leaders using the neuro-evolutionary learning algorithm can influence a robot swarm to significantly outperform a baseline model that does not incorporate learning.

4.1 Experimental Design

The independent variables are the leadership model, swarm size and the leadership percentage. The leadership model is either the learning based model described in Sect. 3, or a baseline model. Baseline leaders do not learn and always align their heading towards the goal. Swarms of 50 and 100 agents were evaluated with leadership percentages ranging from 4% to 24%, as shown in Table 1. These percentages reflect observations on leadership in biological swarms [11,26,30].

The swarm begins each rally trial gathered at a starting point d_{init} distance units (u) from, and at a random angle to the goal. $d_{init} = 400$ distance units (u) to minimize locating the goal by chance, while also completing the trial within a reasonable number of time steps. A zero-mean Gaussian noise with variance σ_{init} was added to the starting positions of each agent, in order to avoid collisions. Swarm agents are initialized with uniform random orientations, and the swarm's starting speed is set to v_{init}, which is 50% of the swarm's maximum speed, v_{max}. This stochasticity encourages the learning of generalized behaviors. The r_{rep}, r_{ori}, and r_{att} radii govern the non-leader agents' motion, and were selected to be 20 u, 30 u, and 50 u, respectively, based on biological swarms [15]. The total number of NNs, m, and the number of parents, λ, were set to be 15 and 5, respectively. The training session lasted 400 epochs, ensuring convergence of all training errors.

The percentage of non-leader agents within a radius of the goal location, r_{goal}, is calculated at trial completion, and averaged over all trials to calculate the percent reached (PR). The test error (E) represents the accumulated distance to the goal, and is calculated using Eq. 1.

Leaders were trained using all independent variable configurations. After 400 training epochs, each NN was evaluated over 100 trials without any mutations, and the NN with the minimum root-mean-squared error deemed the *champion*. The process was repeated 10 times, resulting in 10 champions. The champion NNs' performance metrics are reported in all results.

Table 1. Experimental parameters and independent variables.

Parameters	Values	Parameters	Values
Swarm size	50, 100	r_{rep}	20 u
Leader percentage	4%, 8%, 12%, 16%, 20%, 24%	r_{ori}	30 u
τ	20000 steps	r_{atr}	50 u
v_{init}	1 u/step	r_{goal}	150 u
σ_{init}	50	v_{max}	2 u/step
λ	5	m	15

4.2 Results

The performance, percent reached (PR), improved with increasing leadership percentage, as shown in Table 2 and Fig. 1. The learning based agents successfully guided the swarm, with both 50 and 100 agents, even with the 4% leadership. However, the baseline leaders generally failed to lead any swarm members to the goal with the 4% and 8% leadership. The baseline model matched or exceeded the learning model when leaders composed 20% and 24% of the swarm, but pairwise T-tests (degrees of freedom [dof] = 999 in all tests) found no significant differences between the models. The learning based method significantly outperformed the baseline for all other cases ($p < 0.01$). The PR was generally better with a swarm size of 50 and pairwise T-tests comparing PR by swarm size found significant differences only at the 4% and 8% leadership percentages ($p < 0.01$).

Table 2. The percent reached (PR) descriptive statistics (mean - (μ), standard error - SE) by swarm size, leadership percentage, and leadership model.

Leader %	50 agents				100 agents			
	Baseline		Learning		Baseline		Learning	
	μ	SE	μ	SE	μ	SE	μ	SE
4%	0.04	0.06	32.01	2.75	0.01	0.02	20.31	2.38
8%	0.02	0.04	54.24	3.01	0.01	0.02	40.33	2.94
12%	21.21	8.01	69.73	2.81	5.05	4.29	69.52	2.79
16%	64.67	9.36	84.77	2.21	43.43	9.71	86.34	2.08
20%	89.92	5.89	87.48	2.02	75.76	8.40	90.00	1.82
24%	97.98	2.76	93.71	1.49	93.94	4.68	85.42	2.14

The test error (E) results, presented in Table 3, were grouped into bins (size = 10), as shown in Fig. 2. The E for a majority of trials (>80%) was less than 400, and trials with E $\geq$ 400 were deemed unsuccessful, and are grouped into the final bin. E is impacted by the time required to reach the goal, but there is

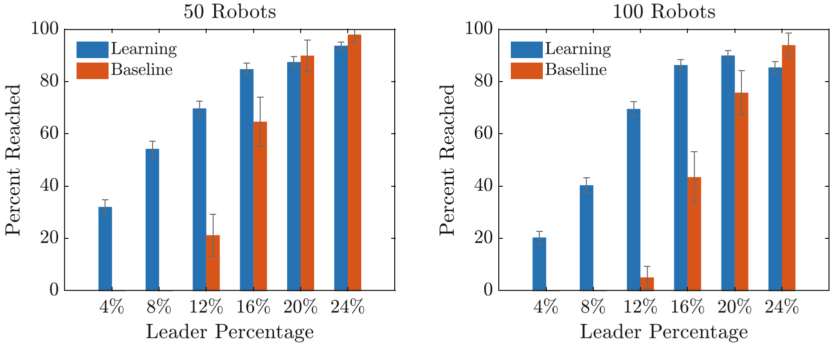

Fig. 1. Percent reached by leadership model, leadership percentage and swarm size.

no effect on the PR if the agents reach the goal by trial completion. Thus, slow moving swarms have higher Es. The PR metric suggests that agents occasionally reach the goal area, even with 4% and 8% leadership, but with higher minimum errors than swarms with higher leadership percentages. The leaders move slower when their percentage is low, and faster when their percentage is higher.

Table 3. The test error (E) descriptive statistics by leadership model, swarm size and leadership percentage ($L\%$). Median is reported due to a large number of outliers.

L%	50 agents						100 agents					
	Baseline			Learning			Baseline			Learning		
	Med	Min	Max	Med	Min	Max	Med	Min	Max	Med	Min	Max
4%	308	261	1033	193	67	20992	313	262	993	376	76	19742
8%	302	255	923	140	38	7361	307	274	766	197	43	20247
12%	282	40	320	81	22	6268	304	42	324	83	32	6236
16%	38	36	313	31	20	6540	271	37	321	30	21	6528
20%	36	35	307	29	19	7721	37	36	317	23	20	2681
24%	35	34	317	21	18	9092	35	34	308	25	19	8893

Median and minimum Es of the learning model were lower than the baseline with leadership percentages higher than 16%, despite the PR not being statistically significant, suggesting that learning agents move faster than the baseline agents. Trials where the fast moving learning leaders fail to guide the swarm explain the high maximum E, as the swarms travel farther away from the goal. The median and minimum Es decreased with increasing leadership percentage. The largest reduction in the median occurred when the percentage increased from 4% to 12%. The change in the median and minimum Es was minimal past the 16% leadership percentage.

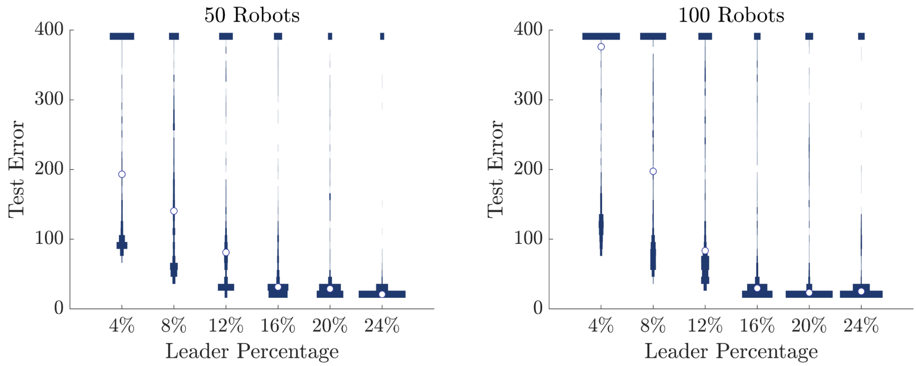

Fig. 2. Learning model test errors by leadership percentage and swarm size. Errors are packed into bins of size 10, and errors greater than 400 are grouped together into the top bin (400). The circles represent the median error.

5 Discussion

The learning model resulted in leaders that successfully influenced the swarm to achieve the rally task with very small leadership percentages, which validates the proposed learning based controller and answers the primary research question. Generally, the learning model leaders outperformed the baseline model and were able to perform significantly better at the lowest leadership percentages. While swarms led by small sets of leaders took longer and were less likely to reach the goal, they were able to achieve the task at leadership percentages representative of biological species, which can be as low as 5% [30] and 9% [26].

Lower leadership percentages (i.e., 4%) require the leaders to learn more nuanced behaviors in order to be effective, which explains the slow convergence of the training errors relative to higher percentages. The leaders' movements are fast, aggressive and goal driven when their influence over the swarm is high, and resemble the baseline model. Thus, the baseline method is only viable if the leader influence is guaranteed to be high throughout the duration of the task.

The biological literature demonstrates that leaders must balance goal-oriented actions with socially-oriented ones in order to be effective [16]. Leaders following the learning model act based on both the goal and their followers, while the baseline leaders are indifferent to their followers. The significant performance differences between the two models emphasize the importance of spatial awareness, and confirms that the learning based model successfully combines goal-oriented and socially-oriented actions.

Biological swarms rely only on local interactions, and typically use an implicit leadership mechanism [9]. The learning based strategy draws inspiration from biological swarms in that it is based on implicit communication and local decision making. Further, no agent knows whether another agent is a leader or not.

The leadership percentage strongly affects the characteristics of the learned behaviors. The learned behaviors are more aggressive, and the leaders travel straight towards the goal when their overall influence is higher. However, the leaders follow more complex movement patterns when the leadership percentage is low. Leaders must be aware of their followers when the leaders' influence is low, and use these complex movement patterns in order to ensure they are being followed, otherwise the leaders lose track of the swarm. The proposed neuro-controller solves this problem by integrating both the follower positions and the goal position into the decision making process, which enables reliable swarm control with only a small percentage of informed leaders.

6 Conclusion

A learning based leadership strategy was developed that allowed small percentages of leaders to drastically improve its task performance over a baseline model. The leadership model incorporates implicit leadership and communication, which allows any agent to assume a leadership role at any given time. While a higher leadership percentage improved task performance, the increase was minimal with percentages >16%. The task was successful with leadership percentages as low as 4%, but the consistency of success increased with higher percentages.

Acknowledgments. This work was partially supported by NSF Grant #1723924 and DARPA award W31P4Q18C0034.

References

1. Ampatzis, C., Tuci, E., Trianni, V., Dorigo, M.: Evolution of signaling in a multi-robot system: categorization and communication. Adapt. Behav. **16**(1), 5–26 (2008)
2. Baldassarre, G., Nolfi, S., Parisi, D.: Evolving mobile robots able to display collective behaviors. Artif. Life **9**(3), 255–267 (2003)
3. Baldassarre, G., Trianni, V., Bonani, M., Mondada, F., Dorigo, M., Nolfi, S.: Self-organized coordinated motion in groups of physically connected robots. IEEE Trans. Syst. Man Cybern. Part B (Cybern.) **37**(1), 224–239 (2007)
4. Barca, J.C., Sekercioglu, Y.A.: Swarm robotics reviewed. Robotica **31**(3), 345–359 (2013)
5. Beekman, M., Fathke, R.L., Seeley, T.D.: How does an informed minority of scouts guide a honeybee swarm as it flies to its new home? Anim. Behav. **71**(1), 161–171 (2006)
6. Beer, R.D., Gallagher, J.C.: Evolving dynamical neural networks for adaptive behavior. Adapt. Behav. **1**(1), 91–122 (1992)
7. Brambilla, M., Ferrante, E., Birattari, M., Dorigo, M.: Swarm robotics: a review from the swarm engineering perspective. Swarm Intell. **7**(1), 1–41 (2013)
8. Cavagna, A., et al.: Scale-free correlations in starling flocks. Proc. Natl. Acad. Sci. **107**(26), 11865–11870 (2010)
9. Colby, M.K., Chung, J.J., Tumer, K.: Implicit adaptive multi-robot coordination in dynamic environments. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 5168–5173 (2015)

10. Couzin, I.D., Krause, J.: Collective memory and spatial sorting in animal groups. Theoret. Biol. **218**, 1–11 (2002)
11. Couzin, I.D., Krause, J., Franks, N.R., Levin, S.A.: Effective leadership and decision-making in animal groups on the move. Nature **433**(7025), 513–516 (2005)
12. Pérez, I.F., Boumaza, A., Charpillet, F.: Learning collaborative foraging in a swarm of robots using embodied evolution. In: The European Conference on Artificial Life, pp. 162–161 (2017)
13. Garnier, S., Gautrais, J., Theraulaz, G.: The biological principles of swarm intelligence. Swarm Intell. **1**(1), 3–31 (2007)
14. Gross, R., Dorigo, M.: Towards group transport by swarms of robots. Int. J. Bio-Inspir. Comput. **1**(1–2), 1–13 (2009)
15. Haque, M., Ren, C., Baker, E., Kirkpatrick, D., Adams, J.A.: Analysis of swarm communication models. In: International Workshop on Combinations of Intelligent Methods and Applications, p. 29 (2016)
16. Ioannou, C.C., Singh, M., Couzin, I.D.: Potential leaders trade off goal-oriented and socially oriented behavior in mobile animal groups. Am. Nat. **186**(2), 284–293 (2015)
17. Kilgour, R., Scott, T.: Leadership in a herd of dairy cows. Proc. N. Z. Soc. Anim. Prod. **19**, 36–43 (1959)
18. King, A.J., Johnson, D.D., Van Vugt, M.: The origins and evolution of leadership. Curr. Biol. **19**(19), R911–R916 (2009)
19. Krause, J., Hoare, D., Krause, S., Hemelrijk, C., Rubenstein, D.: Leadership in fish shoals. Fish Fish. **1**(1), 82–89 (2000)
20. Leca, J.B., Gunst, N., Thierry, B., Petit, O.: Distributed leadership in semifree-ranging white-faced capuchin monkeys. Anim. Behav. **66**(6), 1045–1052 (2003)
21. Meese, G., Ewbank, R.: Exploratory behaviour and leadership in the domesticated pig. Br. Vet. J. **129**(3), 251–259 (1973)
22. Panait, L., Luke, S.: Cooperative multi-agent learning: the state of the art. Auton. Agents Multi-Agent Syst. **11**(3), 387–434 (2005)
23. Pini, G., Tuci, E.: On the design of neuro-controllers for individual and social learning behaviour in autonomous robots: an evolutionary approach. Connect. Sci. **20**(2–3), 211–230 (2008)
24. Portugal, S.J., et al.: Upwash exploitation and downwash avoidance by flap phasing in ibis formation flight. Nature **505**(7483), 399–402 (2014)
25. Pugh, J., Martinoli, A.: Parallel learning in heterogeneous multi-robot swarms. In: IEEE Congress on Evolutionary Computation, pp. 3839–3846 (2007)
26. Reebs, S.G.: Can a minority of informed leaders determine the foraging movements of a fish shoal? Anim. Behav. **59**(2), 403–409 (2000)
27. Reynolds, C.W.: Flocks, herds and schools: a distributed behavioral model. Comput. Graph. **21**(4), 25–34 (1987)
28. Ripon, K.S.N., Jakobsen, E., Tannum, C., Montanier, J.M.: Assessing the effect of self-assembly ports in evolutionary swarm robotics. In: IEEE Symposium Series on Computational Intelligence, pp. 1–8 (2016)
29. Sato, S.: Leadership during actual grazing in a small herd of cattle. Appl. Anim. Ethol. **8**(1–2), 53–65 (1982)
30. Seeley, T.D.: The Wisdom of the Hive: The Social Physiology of Honey Bee Colonies. Harvard University Press, Cambridge (2009)
31. Soysal, O., Bahçeci, E., Şahİn, E.: Aggregation in swarm robotic systems: evolution and probabilistic control. Turk. J. Electr. Eng. Comput. Sci. **15**(2), 199–225 (2007)

32. Trianni, V., Groß, R., Labella, T.H., Şahin, E., Dorigo, M.: Evolving aggregation behaviors in a swarm of robots. In: Banzhaf, W., Ziegler, J., Christaller, T., Dittrich, P., Kim, J.T. (eds.) ECAL 2003. LNCS, vol. 2801, pp. 865–874. Springer, Heidelberg (2003). https://doi.org/10.1007/978-3-540-39432-7_93
33. Wolpert, D.H., Tumer, K.: An introduction to collective intelligence. arXiv preprint arXiv:cs/9908014 (1999)
34. Xu, X.K., Kattas, G.D., Small, M.: Reciprocal relationships in collective flights of homing pigeons. Phys. Rev. E **85**(2), 026120 (2012)