Measuring Resilience in Collective Robotic Algorithms

Extended Abstract

Jennifer Leaf
Oregon State University
Corvallis, Oregon, United States
leafj@oregonstate.edu

Julie A. Adams
Oregon State University
Corvallis, Oregon, United States
julie.a.adams@oregonstate.edu

ABSTRACT

Measuring and comparing resilience is crucial for evaluating different algorithms' performance. Existing resilience metrics focus on a system's ability to maintain a particular state, but are inadequate to evaluate whether a system can achieve a novel state after an unexpected disturbance. The presented *resilience power* metric is used to analyze two best-of-N algorithms. Both algorithms exhibited high resilience power when changing the collective's population size, but this result did not correlate with high overall task performance.

KEYWORDS

resilience; swarm intelligence; collective decision-making

ACM Reference Format:
Jennifer Leaf and Julie A. Adams. 2022. Measuring Resilience in Collective Robotic Algorithms: Extended Abstract. In *Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), Online, May 9–13, 2022*, IFAAMAS, 3 pages.

1 INTRODUCTION

Robotic systems are increasingly operating in uncertain and dynamic environments. Robots must exhibit *resilience*, meaning resistance to unexpected disruptions, or recovery from unavoidable events through adaptive behaviors. Quantitatively comparing resilience between different algorithms is necessary to evaluate robotic systems' performance. Robotic collectives offer unique advantages for certain tasks, such as searching a large spatial area. Real-world operations can introduce unexpected changes to the collective or its operational environment. The collective's algorithm must adjust to current circumstances and recover from unexpected setbacks. Resilience metrics must measure how well an algorithm copes with unexpected events, and enable evaluation of whether alternative algorithms improve the collective's performance.

2 RELATED RESEARCH

The best-of-N consensus decision making problem requires agents select from two or more alternatives (e.g., honeybee nest selection [13, 14]). Scouts search for prospective site locations, and recruit other scouts to consider the discovered options. The best-of-N problem is an achievement oriented task, where the task is complete after a particular goal state is reached [16]. Existing resilience metrics were primarily developed for equilibrium based control systems seeking to maintain a desired property/state over time (e.g., [1, 4–7]). These metrics assess the deviation from a nominal state, and the time and effort required to restore the system's prior performance.

Studies of disturbances' effects on robotic collectives are limited. Agent-based consensus decision-making implementations generally focus on unchanging environmental configurations and collective sizes; however, both the number of agents and the potential sites' spatial distribution can affect the collective's selection accuracy [9, 10, 15]. Understanding external disturbances' impact on the decision process is critical to designing resilient collectives.

3 CONSENSUS ALGORITHMS

The performance and resilience of two best-of-N algorithms are investigated. The A_{core} algorithm implements Reina's agent-based model [11]. Robots probabilistically transition between uncommitted and favoring states while forming a consensus. Robots in latent states explore the environment, while interactive robots wait in the hub. Robots that sense a quorum transition to a committed state and no longer consider alternatives. A limitation of A_{core} is *environmental bias*, where a site closer to the hub is selected over the highest valued site [10, 15]. The collective can build sufficient support for the nearby site and select it before robots favoring the highest valued site can recruit enough uncommitted robots [8, 12].

A_{ext} extends A_{core} in two ways. A_{ext} introduces state transition delays that enable simultaneous comparison of sites. Robots favoring nearby sites are delayed more than robots favoring distant sites. A_{ext} also increases the interaction rate of robots favoring more distant sites relative to each sites' distance to the hub [2, 3].

Disturbances affect the outcome by altering the states, the possible state transitions induced by an agent's actions, the state transition probabilities, or a combination thereof. Disturbances can alter the outcome even if a robot or the collective is not directly affected.

4 EXPERIMENTAL DESIGN

The hypothesis was that A_{ext} will achieve greater or equal selection accuracy as A_{core} when agents are removed during runtime (*Remove Agents* disturbance). The task successfully completed if the highest valued site was selected, and failed if any other site was selected or there was no consensus after 100,000 iterations. Fifty trials were conducted for each combination of independent variables and site configurations, as well as with no disturbance. The **independent variables** (see Table 1) were the number of agents at a trial's start, the site locations and qualities (listed in increasing distance from the hub), the disturbance timing relative to the decision progress, and the change in collective population size. The **dependent variables** were selection accuracy and resilience power. *Selection accuracy* represents the percentage of trials

Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), P. Faliszewski, V. Mascardi, C. Pelachaud, M.E. Taylor (eds.), May 9–13, 2022, Online.

Table 1: Independent variables

Variable	Values
Number of Agents	100, 300, 500
Site Value	60, 70, 80, 90
Site Locations	250-400m from the hub
Disturbance Timing	15%, 35%, 60%
Agents Removed	25%, 50%
Site configurations	
SC.1	Easy (90, 80, 70, 60)
SC.2	Easy (60, 90, 80, 70)
SC.3	Hard (80, 70, 60, 90)

in which the collective selected the best site. The resilience *power* is the probability that an algorithm starting at a state s_0 successfully reaches a goal state before a time deadline T:

$$\theta(T; s_0) = \Big(P_G(\mathtt{true}|s_0; T)\Big).$$

This metric permits comparing an algorithm's real-world performance to its performance in ideal conditions.

5 RESULTS

Selection accuracy for both algorithms after the Remove Agents disturbance demonstrated similar trends to their respective baselines. A_{core}'s baseline performance ranged from 90-100% for the SC.1 easy site configuration, 60-94% for the SC.2 easy configuration, and 0-8% for SC.3, shown as A_{core} in Figure 1. A_{core}'s post-disturbance selection accuracy ranged from 82% (original collective = 100 agents) to 98-100% (500 agents) for SC.1, 66-70% (100 agents) to 94-96% (500 agents) for SC.2, and 12% (100 agents) to 0% (500 agents) for SC.3.

A_{ext}'s baseline selection accuracy ranged from 72-100% for SC.1, 62-90% for SC.2, and 52-88% for SC.3, shown as A_{ext} in Figure 1. The selection accuracy after the disturbance ranged from 68% (original collective = 100 agents) to 94-96% (500 agents) for SC.1, 72% (100 agents) to 88-92% (500 agents) for SC.2, and 58% (100 agents) to 80-86% (500 agents) for SC.3.

A_{core}'s resilience power varied less than A_{ext}, from 7 percentage points below the baseline to 0 points above (SC.1), and increased to 1-7 points above the baseline (SC.2). SC.3's power was 0-3 percentage points above the baseline.

A_{ext}'s resilience power for SC.1 was 4 percentage points below the baseline for all starting population sizes. SC.2's power was between 3 points below the baseline (300 agents) to 5 points above (100 agents). Finally, SC.3 had the lowest power with 500 agents (2 points below the baseline), while collectives with 100 agents had a power measurement 4 points above the baseline.

6 DISCUSSION

The hypothesis that the A_{ext} algorithm, with its extensions to mitigate environmental bias, will demonstrate equal or better performance than the A_{core} algorithm was supported. Changing the collective's population size at runtime did not cause a substantial change in either algorithm's performance relative to the baseline. The resilience power metric shows that both algorithms are *resilient*

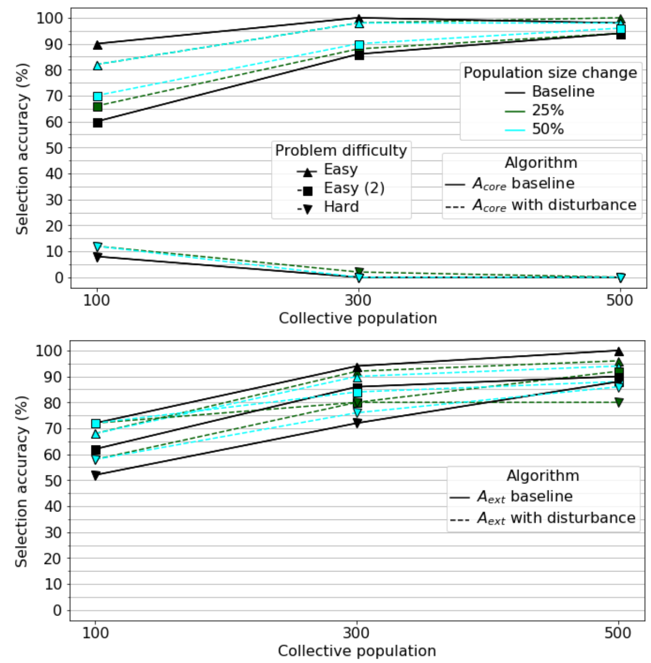

Figure 1: Results for the Remove Agents disturbance by original collective size, population size change, and problem difficulty. The results for each algorithm are plotted separately.

to the Remove Agents disturbance, due to the relatively small difference in power between the baseline and post-disturbance results. However, these results also illustrate that *a resilient algorithm is not necessarily a high-performing algorithm.* A_{ext}'s much greater selection accuracy for the hard problem (SC.3) demonstrates that optimizing for resilience alone is an insufficient goal. Resilience is not a useful attribute if the algorithm has a poor success rate.

The collective's selection accuracy after the disturbance was similar to the undisturbed baseline, because the agents' state transition rates were unaffected. The disturbance did not directly alter the percentage of agents favoring a particular site over a different site, because agents were removed without regard to their current state (i.e., the relative proportions of agents in each state was preserved). The disturbance's timing also did not affect the collective's selection accuracy, because the timing did not affect the rates of initiating recruitment or inhibition interactions. A_{ext}'s higher performance in the hard site configuration demonstrated that the algorithm extensions mitigated environmental bias, even when the collective's population size changed during runtime.

The resilience power metric provides a comparison of the algorithms' performance between baseline and disturbance configurations. The calculation of the power metric was independent of the algorithms or disturbance variables, which demonstrates generalizability to other disturbance types and collective tasks.

ACKNOWLEDGMENTS

This work was supported by ONR Awards N00014-16-1-3025 and N00014-18-1-2831, and an ARCS Foundation Scholar award (Leaf).

REFERENCES

[1] Michel Bruneau, Stephanie E Chang, Ronald T Eguchi, George C Lee, Thomas D O'Rourke, Andrei M Reinhorn, Masanobu Shinozuka, Kathleen Tierney, William A Wallace, and Detlof Von Winterfeldt. 2003. A Framework to Quantitatively Assess and Enhance the Seismic Resilience of Communities. *Earthquake Spectra* 19, 4 (2003), 733–752.

[2] Jason R. Cody. 2018. *Discrete Consensus Decisions in Human-Collective Teams*. Ph.D. Dissertation. Vanderbilt University.

[3] Jason R. Cody, Karina A. Roundtree, and Julie A. Adams. 2021. Human-collective collaborative target selection. *ACM Transactions on Human-Robot Interaction* 10, 2 (2021), 1–29.

[4] Devanandham Henry and Jose Emmanuel Ramirez-Marquez. 2012. Generic metrics and quantitative approaches for system resilience as a function of time. *Reliability Engineering and System Safety* 99 (2012), 114–122.

[5] Dave Hodgson, Jenni L. McDonald, and David J. Hosken. 2015. What do you mean, 'resilient'? *Trends in Ecology & Evolution* 30, 9 (2015), 503–506.

[6] Seyedmohsen Hosseini, Kash Barker, and Jose E. Ramirez-Marquez. 2016. A review of definitions and measures of system resilience. *Reliability Engineering and System Safety* 145 (2016), 47–61.

[7] Johannes Ingrisch and Michael Bahn. 2018. Towards a Comparable Quantification of Resilience. *Trends in Ecology & Evolution* 33, 4 (2018), 251–259.

[8] Teeraphan Laomettachit, Teerasit Termsaithong, Anuwat Sae-Tang, and Orawan Duangphakdee. 2015. Decision-making in honeybee swarms based on quality and distance information of candidate nest sites. *Journal of Theoretical Biology* 364 (2015), 21–30.

[9] Jennifer Leaf and Julie A. Adams. 2021. The Impact of Disturbances on Consensus Decision Making in Collective Agent Systems. (2021). The 4th International Symposium on Swarm Behavior and Bio-Inspired Robotics (SWARM).

[10] Andreagiovanni Reina, Roman Miletitch, Marco Dorigo, and Vito Trianni. 2015. A quantitative micro-macro link for collective decisions: the shortest path discovery/selection example. *Swarm Intelligence* 9, 2-3 (2015), 75–102.

[11] Andreagiovanni Reina, Gabriele Valentini, Cristian Fernández-Oto, Marco Dorigo, and Vito Trianni. 2015. A Design Pattern for Decentralised Decision Making. *PLOS ONE* 10, 10 (2015), e0140950.

[12] T. M. Schaerf, J. C. Makinson, M. R. Myerscough, and M. Beekman. 2013. Do small swarms have an advantage when house hunting? The effect of swarm size on nest-site selection by Apis mellifera. *Journal of the Royal Society Interface* 10 (2013), 20130533.

[13] Thomas D Seeley. 2010. *Honeybee democracy*. Princeton University Press, Princeton, NJ.

[14] Thomas D. Seeley, P. Kirk Visscher, Thomas Schlegel, Patrick M. Hogan, Nigel R. Franks, James A. R. Marshall, and Honeybee. 2012. Stop Signals Provide Cross Inhibition in Collective Decision-Making by Honeybee Swarms. *Science* 335, January (2012), 108–111.

[15] Gabriele Valentini, Eliseo Ferrante, and Marco Dorigo. 2017. The Best-of-n Problem in Robot Swarms: Formalization, State of the Art, and Novel Perspectives. *Frontiers in Robotics and AI* 4 (2017).

[16] M. Birna Van Riemsdijk, Mehdi Dastani, and Michael Winikoff. 2008. Goals in agent systems: A unifying framework. In *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS*, Vol. 2. 702–709.