

Supplementary Material for Deep Reinforcement Learning for Decentralized Multi-Robot Exploration with Macro Actions

Aaron Hao Tan, *Student Member, IEEE*, Federico Pizarro Bejarano, Yuhan Zhu, Richard Ren and Goldie Nejat, *Member, IEEE*

A. Comparison Study: Statistical Analysis

A Friedman test ($n = 6, k = 40, \alpha = 0.05, \chi^2(5)$) was conducted for the metrics: 1) total number of steps (TNS), b) total travel distance (TTD), c) number of local interactions (NOLI), and d) objective function values (OFV) across the CSPs of 100%, 80%, 50% and 0%. The results showed a statistically significant difference ($p < 0.0001$) between the performance metrics across all communication success probability (CSP), top of Table I. Using post-hoc Wilcoxon Signed-rank tests with a Bonferroni correction ($\alpha_{revised} = 0.001$), a statistically significant difference ($p < 0.001$) was also found for each of the four performance metric between MADE-Net and each benchmark method (color coded), bottom of Table I. MADE-Net had an average variance of 2.5, 4.59, 0.01, and 1.03 for TNS, TTD, NOLI and OFV across the CSPs, compared to 3.41, 66.93, 2.5, and 1.68 for MADE-Net-DT, respectively. Thus, the trend in the variances for MADE-Net-DT across each CSP shows that it is more affected by communication dropouts compared to MADE-Net.

TABLE A.I. FRIEDMAN TEST AND WILCOXON SIGNED-RANK TEST FOR COMPARISON STUDY

Friedman Test				
Metric	$\chi^2(5)$ at each CSP where all $p < 0.0001$			
	100%	80%	50%	0%
Average TNS	88.92	63.61	57.98	82.42
Average TTD	147.03	102.04	115.94	100.02
Average NOLI	165.14	144.19	152.82	155.14
Average OFV	417.68	411.89	363.18	466.77
Wilcoxon Signed-Rank Test with Pairwise Comparison				
Metric	Z Statistic between MADE-Net and (<i>MADE-Net-DT</i> , <i>DME-DRL</i> , <i>PB</i> , <i>UB</i> , <i>NF</i>) at each CSP			
	100%	80%	50%	0%
Average TNS $p < 0.001$	-5.4303, -5.2374, -4.2273, -4.6372, -4.5297	-4.8703, -5.7821, -3.2595, -3.9450, -3.5149	-4.8005, -4.2846, -3.6004, -4.1668, -3.8644	-5.1413, -5.1989, -2.7556, -3.0377, -3.0631
Average TTD $p < 0.001$	-5.0270, -5.4999, -5.4840, -5.5109, -5.4975	-4.5297, -5.5101, -5.1480, -5.3362, -5.4303	-5.4034, -5.0082, -5.2152, -5.5109, -5.4572	-5.3496, -5.0257, -5.0136, -5.2690, -5.5109
Average NOLI $p < 0.001$	-5.1923, -6.3639, -5.5359, -5.4921, -5.7383	-4.2998, -5.5442, -5.4168, -5.5109, -5.4975	-4.0769, -5.8869, -5.5222, -4.9892, -5.3332	-3.5584, -5.1854, -5.3731, -5.2894, -5.5539
Average OFV $p < 0.001$	-6.8044, -6.9305, -8.6818, -8.6783, -8.2234	-7.5402, -7.5109, -8.5202, -8.6818, -8.6214	-6.2474, -8.9409, -8.5099, -8.5993, -8.6783	-6.4572, -7.9806, -8.2955, -8.3889, -8.3459

B. Scalability Study: Statistical Analysis

A Friedman test ($n = 5, k = 12, \alpha = 0.05, \chi^2(4)$) was conducted on average exploration time for each environment ($20 \times 20 \text{ m}^2$, $30 \times 30 \text{ m}^2$, $40 \times 40 \text{ m}^2$, and $60 \times 60 \text{ m}^2$) and team size (teams of 2 to 6 mobile robots). A statistically significant difference ($p < 0.05$) was found between the exploration time of all team sizes (2 to 6) across all environment sizes ($20 \times 20 \text{ m}^2$, $40 \times 40 \text{ m}^2$, $30 \times 30 \text{ m}^2$, and $60 \times 60 \text{ m}^2$), top of Table B.I. Pairwise post-hoc analysis using Wilcoxon Signed-Rank tests were conducted for each robot team size combination, bottom of Table B.I. The results showed there were no statistically significant difference ($p > 0.05$) for team sizes of 4 and 5 robots, 5 and 6 robots, and 4 and 6 robots in environment sizes of $20 \times 20 \text{ m}^2$, $40 \times 40 \text{ m}^2$, and $60 \times 60 \text{ m}^2$. For the $30 \times 30 \text{ m}^2$ environment, no statistical difference was found only for team sizes of 5 and 6 robots ($p > 0.05$).

TABLE B.I. FRIEDMAN TEST AND WILCOXON SIGNED-RANK TEST FOR SCALABILITY STUDY

Friedman Test				
Team Size Groups	$(\chi^2(4), p)$ for each team size groups' exploration time in each env. size			
	$20 \times 20 \text{ m}^2$	$30 \times 30 \text{ m}^2$	$40 \times 40 \text{ m}^2$	$60 \times 60 \text{ m}^2$
(2, 3, 4, 5, 6)	(4.67, 0.01932)	(10.93, 0.02732)	(11.47, 0.02179)	(8.27, 0.02433)
Wilcoxon Signed-Rank Test with Pairwise Comparison				
Team Size Pairs	(Z, p) for each team size pairs' exploration time in each env. size			
	$20 \times 20 \text{ m}^2$	$30 \times 30 \text{ m}^2$	$40 \times 40 \text{ m}^2$	$60 \times 60 \text{ m}^2$
(2, 3)	(-1.033, 0.00124)	(-1.069, 0.00254)	(-1.389, 0.00209)	(-2.069, 0.00466)
(2, 4)	(-2.353, 0.00329)	(-1.396, 0.00598)	(-2.843, 0.00349)	(-1.204, 0.00998)
(2, 5)	(-3.566, 0.00738)	(-1.289, 0.00183)	(-4.882, 0.00395)	(-2.543, 0.00333)
(2, 6)	(-4.921, 0.00829)	(-2.642, 0.00582)	(-3.692, 0.00193)	(-4.845, 0.00984)
(3, 4)	(-5.322, 0.00182)	(-3.221, 0.00693)	(-1.604, 0.00826)	(-0.534, 0.00258)
(3, 5)	(-3.221, 0.00582)	(-3.686, 0.00839)	(-2.726, 0.00922)	(-3.529, 0.00345)
(3, 6)	(-7.893, 0.00284)	(-3.503, 0.00598)	(-3.604, 0.00589)	(-4.488, 0.00199)
(4, 5)	(-2.503, 0.05922)	(-3.849, 0.00395)	(-1.604, 0.05822)	(-2.604, 0.05251)
(4, 6)	(-3.543, 0.05201)	(-4.839, 0.00912)	(-2.842, 0.07501)	(-3.572, 0.09251)
(5, 6)	(-3.799, 0.06836)	(-1.403, 0.54142)	(-1.292, 0.25034)	(-0.535, 0.75083)